UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH

Threshold phenomena involving the connected components of random graphs and digraphs

Matthew Coulson

Threshold phenomena involving the connected components of random graphs and digraphs

Matthew Coulson

Thesis advisor:

Guillem Perarnau

Abstract

We consider some models of random graphs and directed graphs and investigate their behaviour near thresholds for the appearance of certain types of connected components.

Firstly, we look at the critical window for the appearance of a giant strongly connected component in binomial random digraphs. We provide bounds on the probability that the largest strongly connected component is very large or very small.

Next, we study the configuration model for graphs and show new upper bounds on the size of the largest connected component in the subcritical and barely subcritical regimes. We also show that these bounds are tight in some instances.

Finally we look at the configuration model for random digraphs. We investigate the barely sub-critical region and show that this model behaves similarly to the binomial random digraph whose barely sub- and super- critical behaviour was studied by Łuczak and Seierstad. Moreover, we show the existence of a threshold for the existence of a giant weak component, as predicted by Kryven.

Resum

En aquesta tesi considerem diversos models de grafs i graf dirigits aleatoris, i investiguem el seu comportament a prop dels llindars per l'aparició de certs tipus de components connexes.

En primer lloc, estudiem la finestra crítica per a l'aparició d'una component fortament connexa en dígrafs aleatoris binomials (o d'Erdős-Rényi). En particular, provem diversos resultats sobre la probabilitat límit que la component fortament connexa sigui sigui molt gran o molt petita.

A continuació, estudiem el model de configuració per a grafs no dirigits i mostrem noves cotes superiors per la mida de la component connexa més gran en els règims subcrítics i quasi-subcrítics. També demostrem que, en general, aquestes cotes no poden ser millorades.

Finalment, estudiem el model de configuració per a dígrafs aleatoris. Ens centrem en la regió quasi-subcrítica i demostrem que aquest model es comporta de manera similar al model binomial, el comportament del qual va ser estudiat per Łuczak i Seierstad en les regions quasi-subcrítica i quasi-supercrítica. A més a més, demostrem l'existència d'una funció llindar per a l'existència d'una component feble gegant, tal com va predir Kryven.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my supervisor, Guillem Perarnau. He has always been encouraging and enthusiastic about all of the problems which I have worked on. I could not have wished for a better supervisor for these last few years.

Furthermore, I would like to thank all of the people who I have worked with on combinatorics. Everyone has a different perspective and it has been incredibly useful to work with different people and learn from them.

This work was supported by the Spanish Ministerio de Economá y Competitividad through the project MTM2017-82166-P and the FPI-scholarship PRE2018-083621.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 Probabilistic Combinatorics and Thresholds

Probabilistic combinatorics is a relatively young area of maths which can trace its roots to the work of Paul Erdős and co-authors approximately 70 years ago. Since then, the area has developed in two different directions. The first of which involves existence proofs utilising the probabilistic method whereby we use the simple fact that if a randomly chosen object has a property with positive probability, then surely there must be an object with this property. This is particularly useful as it is often incredibly difficult to come up with explicit constructions of objects with a given property. A good example of this is Erdős' lower bound on the diagonal Ramsey number $R(s,s) \geq 2^{s/2}$ [23] which was proved with the probabilistic method and is much better than all explicit constructions, none of which are even exponential in s.

The second direction involves properties of random combinatorial objects. Here we sample a combinatorial object from some distribution and any parameter which one may be interested in is now a random variable about which there are many natural questions such as what is its expectation, variance etc. It is this direction which we shall follow in this thesis and we will interest ourselves with thresholds in random graphs and directed graphs.

If $X = X_n$ is a random combinatorial object with parameter, $p(X) = p_n(X_n)$ then we

say that $r = r(n)$ is a *threshold function* for the property $\mathcal{P}$ if the following holds,

$$\mathbb{P}(X_n \in \mathcal{P}) \to 0 \text{ as } n \to \infty \qquad \text{if } p_n(X_n) \ll r(n)$$

$$\mathbb{P}(X_n \in \mathcal{P}) \to 1 \text{ as } n \to \infty \qquad \text{if } p_n(X_n) \gg r(n)$$

Moreover, we will say that r is a *sharp threshold* if for any $\varepsilon > 0$,

$$\mathbb{P}(X_n \in \mathcal{P}) \to 0 \text{ as } n \to \infty \qquad \text{if } p_n(X_n) \leq (1-\varepsilon)r(n)$$

$$\mathbb{P}(X_n \in \mathcal{P}) \to 1 \text{ as } n \to \infty \qquad \text{if } p_n(X_n) \geq (1+\varepsilon)r(n)$$

In both of the above, we have statements of the form $\mathbb{P}(X_n \in \mathcal{P}) \to 1$ as $n \to \infty$. At some points in the remainder of this thesis we shall refer to this as X_n having $\mathcal{P}$ *asymptotically almost surely* (abbreviated as a.a.s.)

In this thesis we look at sharp thresholds for the existence of connected components of linear size in various graph and directed graph models. We also look very closely in at these thresholds and investigate the behaviour of the size of the largest connected component within a $(1 + o(1))$ factor of these thresholds. In the remainder of the chapter we shall introduce the relevant random models and discuss the prior results on the behaviour of the size of the largest component near such thresholds.

1.2 The Erdős Rényi Random Graph

The Erdős Rényi model is the oldest random graph model and we shall denote it by $G(n, p)$. It is a graph on vertex set $[n] = \{1, 2, \ldots, n\}$ formed by including each of the $\binom{n}{2}$ possible edges independently with probability p. Note that this is not the same model originally introduced by Erdős and Rényi who considered the model $G(n, m)$, where we take a uniformly random graph on n vertices and m edges. The two models are however essentially equivalent when one takes $p = m/\binom{n}{2}$ and in particular if $\mathcal{P}$ is any monotone

property, then $\lim_{n\to\infty} \mathbb{P}(G(n,m) \in \mathcal{P}) = \lim_{n\to\infty} \mathbb{P}(G(n,p) \in \mathcal{P})$ (see [44, Section 1.4] for example). As such, it is almost always preferable to work with $G(n,p)$ due to its desirable properties such as the independent edges and we shall phrase all result in this section for $G(n,p)$ however they all hold in $G(n,m)$ with $m = \binom{n}{2}p$.

The component structure of $G(n,p)$ has been of interest since the model was first introduced. In one of the first papers written on the model Erdős and Rényi [24] proved the following result fully describing the component structure for almost all p.

Theorem 1.2.1. Let $c > 0$ be constant, $n \in \mathbb{N}$ and consider the model $G(n,p)$ with $p = c/n$. Then a.a.s.,

- If $c < 1$, the largest component is a tree of size $f(c,n) = \frac{1+o(1)}{c-1-\log(c)}\left(\log(n) - \frac{5}{2}\log\log(n)\right)$;
- If $c = 1$, the largest component is of order $n^{2/3}$;
- If $c > 1$, then the largest component has size $g(c)n$ where $g(c)$ is the unique solution $x > 0$ of $1 - x = e^{-cx}$, and all other components are of order at most $f(c', n')$ for some $c' < 1$ and $n' = (1 - g(c))n$.

The range $c < 1$ is known as the subcritical regime while $c > 1$ is referred to as the supercritical regime. The behaviour of the size of the largest component as c passes 1 going from subcritical to supercritical is known as the "double jump" and is perhaps one of the most surprising properties of random graphs.

Following on from the work of Erdős and Rényi it is natural to ask what happens if $c \to 1$ rather than $c = 1$. The study of this range was started by Bollobas [7] and Łuczak [52] who proved the following,

Theorem 1.2.2. Let $n \in \mathbb{N}$ and take p such that $np = 1 + \varepsilon$, where $\varepsilon = \varepsilon(n) \to 0$, and define $k_0 = 2\varepsilon^{-2}\log(n|\varepsilon|^3)$.

- If $n\varepsilon^3 \to -\infty$ then a.a.s. $G(n,p)$ contains no component of size greater than k_0.

- If $n\varepsilon^3 \to \infty$ then a.a.s. $G(n,p)$ contains a unique component of size greater than k_0. This component has size $2\varepsilon n(1+o(1))$.

When $\varepsilon < 0$ in the above theorem we are in the barely subcritical region and $\varepsilon > 0$ is known as the barely supercritical regime. This only leaves the case where $|\varepsilon| = O(n^{-1/3})$. Note that if we take $\varepsilon \sim n^{-1/3}$ both bounds on the size of the largest component from Theorem 1.2.2 are of order $n^{2/3}$ so by monotonicity it should be unsurprising that the order of the largest component is $n^{2/3}$ for this entire range. We call this range the critical window and parameterise it as $p = n^{-1} + \lambda n^{-4/3}$ for $\lambda \in \mathbb{R}$. The size of the largest component inside the critical window has the property that it is not strongly concentrated as for all other ranges of p. In fact there is a continuous random variable $X = X(\lambda)$ whose domain is all positive reals such that the size of the largest component of $G(n,p)$ is distributed as $Xn^{2/3}$. The exact distribution of X was determined by Pittel [69] and explicit bounds on its tails were given by Nachmias and Peres [63]. Finally, the seminal paper of Aldous [2] shows that the rescaled component sizes behave like the ordered excursion lengths of a Brownian motion with parabolic drift.

Together all of these results provide a thorough description of the component sizes of $G(n,p)$. Furthermore, it provides a picture of what may be true in other random graph models and thus far this picture appears to be correct. Other models appear to go from a subcritical phase through barely subcritical, critical window, barely supercritical and supercritical phases upon the correct choice of parameterisation. We describe how this phase transition happens and our new results in some random graphs and directed graphs in the following sections.

1.2.1 A digraph analogue: $D(n,p)$

The model $D(n,p)$ also known as the *binomial random digraph* is a digraph analogue of $G(n,p)$. It is a digraph with vertex set $[n]$ formed by including each possible directed edge with probability p independently. Note that it is entirely possible to have edges in both

directions between a pair of vertices and this happens with probability p^2 for any given pair. In digraphs there is more than one notion of connected component, in particular there are the following 4 types.

Definition 1.2.3. Let G be a digraph and v a vertex of G then,

- the *weak component* containing v is the component containing v in the underlying undirected graph;
- the *out-component* of v is the set of all vertices u such that there is a path from v to u in G;
- the *in-component* of v is the set of all vertices u such that there is a path from u to v in G;
- the *strong component* of v is the set of all vertices u such that there is both a path from v to u and a path from u to v in G.

Due to the nature of $D(n,p)$ it is simple to reduce the question of the weak, out and in-components to studying components in $G(n,p)$. Thus in $D(n,p)$ the main topic of study is the strong component structure. The strong component structure in $D(n,p)$ is surprisingly close to the component structure of $G(n,p)$ which further reinforces the fact that $D(n,p)$ is a digraph analogue of this model.

The first results on $D(n,p)$ were obtained independently by Karp [47] and Łuczak [53] who showed

Theorem 1.2.4. Let $c > 0$ be constant, $n \in \mathbb{N}$ and define $p = c/n$. Then a.a.s.,

- If $c < 1$, the largest strong component of $D(n,p)$ is a cycle of length $O(1)$;
- If $c > 1$ then the largest strong component of $D(n,p)$ has size $g(c)^2 n$ where $g(c)$ is as defined in Theorem 1.2.1 and all other components are cycles of length $O(1)$.

Note that even the fraction of vertices in the largest component when the graph is supercritical is similar to what is seen in $G(n,p)$. The supercritical regime has particularly

striking similarities - the constant is the square of what is seen in $G(n,p)$. Intuitively this is because one needs both a large in-component and a large out-component, each of which happens with probability approximately $g(c)$. Following on from this result, Łuczak and Seierstad [55] proved the following analogue of Theorem 1.2.2. In this theorem a strongly connected digraph is *complex* if it has more edges than vertices (so it contains multiple cycles).

Theorem 1.2.5. Let $n \in \mathbb{N}$ and $\varepsilon = \varepsilon(n)$ such that $\varepsilon \to 0$ as $n \to \infty$. Choose p so that $np = 1 + \varepsilon$ then,

- If $n\varepsilon^3 \to -\infty$ then a.a.s. every strong component of $D(n,p)$ is an isolated vertex or a cycle of length $O(1/|\varepsilon|)$.
- If $n\varepsilon^3 \to \infty$ then a.a.s. $D(n,p)$ contains a unique complex component of size $4\varepsilon^2 n(1+o(1))$ and every other component is an isolated vertex or a cycle of length $O(1/\varepsilon)$.

This leaves the study of the critical window which again can be found when $p = n^{-1} + \lambda n^{-4/3}$. In this regime, Goldschmidt and Stephenson [31] recently proved an Aldous-type result by giving a scaling limit for the sizes of the largest components. This result in addition to the scaling limit shows that the size of the largest strongly connected component is in fact not concentrated and behaves like $Xn^{1/3}$ for a continuous random variable $X = X(\lambda)$ which can take values on all of the positive real numbers. My contribution to the study of $D(n,p)$ was to prove bounds on the tails of this random variable X. In particular, I proved the following results,

Theorem 1.2.6 (Lower Bound)**.** Let $0 < \delta < 1/800$, $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$. Let $\mathcal{C}_1$ be the largest strong component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then if n is sufficiently large with respect to δ, λ,

$$\mathbb{P}(|\mathcal{C}_1| < \delta n^{1/3}) \leq 2e\delta^{1/4}, \tag{1.1}$$

provided that $\delta \leq \frac{(\log 2)^2}{4|\lambda|^2}$.

Note that the constants in the above theorem have been chosen for simplicity and it is possible to give a bound depending on both λ and δ which imposes no restriction on their relation to one another.

Theorem 1.2.7 (Upper Bound)**.** There exist constants, $\zeta, \eta > 0$ such that for any $A > 0, \lambda \in \mathbb{R}$ the following holds. Let $\mathcal{C}_1$ be the largest strong component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then provided n is sufficiently large with respect to A, λ,

$$\mathbb{P}(|\mathcal{C}_1| > An^{1/3}) \leq \zeta e^{-\eta A^{3/2} + \lambda^+ A}$$

where $\lambda^+ = \max(\lambda, 0)$.

The main idea in the proof of Theorem 1.2.6 is to apply Janson's inequality in order to show there is no cycle of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$ with probability which is at most the bound given in (1.1). This immediately yields Theorem 1.2.6 as a cycle is strongly connected. The proof of Theorem 1.2.7 involves running an exploration process from a large strongly connected subdigraph of $D(n,p)$ to determine whether or not it is a strongly connected component. This requires very good control of certain probabilities and as such we adapt an enumeration of strongly connected digraphs by Pérez-Giménez and Wormald [66] to count such digraphs.

The proofs of these theorems can be found in Chapter 2. These result have been published in [15].

1.3 The configuration model

The configuration model was introduced by Bollobás [6]. It is a model commonly used to pick a random graph with a given degree sequence due to the simplicity of sampling from it and the desirable property that conditionally on the generated graph being simple, it is uniformly random among all graphs with this degree sequence which is often the probability space one wants to study. Furthermore, Janson [41] proved that if the second

moment of the degree sequence is not too large, the configuration model will produce a simple graph with non-negligible probability. This is incredibly useful as it provides a method of transference from the configuration model to the probability space where we pick a graph with given degree sequence uniformly at random.

1.3.1 The configuration model for graphs

In this section, $\mathbf{d}_n = (d_1, \ldots, d_n)$ will be a degree sequence on n vertices and $m = \frac{1}{2}\sum_{i=1}^{n} d_i$ edges and D_n will be the random variable obtained by picking an element of $\mathbf{d}_n$ uniformly at random.

The *configuration model* with degree sequence $\mathbf{d}_n$ denoted $\mathbb{CM}(\mathbf{d}_n)$ is formed by taking d_i "stubs" associated with the vertex i for $i = 1, \ldots, n$, choosing a perfect matching of all $2m$ stubs uniformly at random to form the edges of $\mathbb{CM}(\mathbf{d}_n)$ and contract the stubs associated with vertex i for each i to form the vertices.

It is a natural question to ask what degree sequences have a giant component and whether there is an easy way to describe the threshold (in terms of the degree sequence) above which we find a giant component. Molloy and Reed [58] determined such a threshold in terms of the first two moments of D_n. Define

$$Q = Q(\mathbf{d}_n) = \frac{1}{m}\sum_{i=1}^{n} d_i(d_i - 2).$$

They showed that Q governs whether $\mathbb{CM}(\mathbf{d}_n)$ has a giant component,

Theorem 1.3.1. Let $\mathbf{d}_n$ be a well behaved degree sequence, then

- If $Q(\mathbf{d}_n) < 0$ and $\max \mathbf{d}_n = \Delta_n$, then there exists a constant A such that a.a.s. $\mathbb{CM}(\mathbf{d}_n)$ has no component of size greater than $A\Delta^2 \log(n)$.

- If $Q(\mathbf{d}_n) > 0$, then there exist constants ζ, γ such that a.a.s. $\mathbb{CM}(\mathbf{d}_n)$ has a component of size at least ζn and all other components have size at most $\gamma \log(n)$.

The definition of "well behaved" is somewhat complex and we shall not go into it here. This result has subsequently been improved by other authors, for example [8, 43, 45].

If we consider degree sequences in which $Q = o(1)$, then we cannot apply the result of Molloy and Reed. Hatami and Molloy [35] located the critical window and the order of the largest component inside of it subject to some weak conditions on the degree sequence. Also, inside the critical window Dhara et. al. [17] showed an Aldous-type result under the assumption $\mathbb{E}(D_n^3) = O(1)$. The same authors in a subsequent paper [18] also showed that without this assumption it is possible for the configuration model to fall into a different universality class.

Van der Hofstad, Janson and Luczak [37] considered the barely supercritical case when $Q > 0$, $Q = o(1)$ and we do not lie in the critical window. In this case they were able to show that the size of the largest component is strongly concentrated around a value related to the survival probability of a Galton Watson process with offspring distribution $\hat{D}_n - 1$, where $\hat{D}_n$ is the size biased distribution of D_n. We provide a complementary result to this focussing on the barely subcritical regime, where we instead have $Q < 0$. We were able to show that the components have size bounded above by the maximum possible extinction time of a similar subcritical Galton Watson process as well as computing this extinction time in some cases. Define

$$R = R(\mathbf{d}_n) = \frac{1}{m} \sum_{i=1}^{n} d_i(d_i - 2)^2.$$

We showed the following a.a.s. upper bound on the size of the largest component in the barely subcritical regime.

Theorem 1.3.2. Let $\epsilon > 0$. Let $\mathbf{d}_n$ be a degree sequence that satisfies $\Delta|Q| = o(R)$ and $\mathbb{E}(D_n^4) \leq \Delta^{1/2}$. If $Q \leq -\omega(n)n^{-1/3}R^{2/3}$ for some $\omega(n) \to \infty$, then

$$\mathbb{P}\left(L_1(\mathbb{CM}(\mathbf{d}_n)) \leq (1+\epsilon)\frac{2R}{Q^2}\log\left(\frac{|Q|^3 n}{R^2}\right)\right) = 1 - o(1)\ . \tag{1.2}$$

Note that we can transfer this result to the model in which we pick a random simple

graph with degree sequence $\mathbf{d}_n$ uniformly at random by applying the aforementioned result of Janson [41]. We prove this result by coupling the exploration process with a Galton Watson process and maintaining very tight control on the extinction probability. This control uses exponential tilting in combination with a new local limit theorem. A full proof can be found in Chapter 3. This result is joint work with Guillem Perarnau.

1.3.2 The configuration model for digraphs

In this section, $\vec{\mathbf{d}}_n = (\mathbf{d}_n^-, \mathbf{d}_n^+) = ((d_1^-, d_1^+), \ldots, (d_n^-, d_n^+))$ will be a directed degree sequence on n vertices and $m = \sum_{i=1}^n d_i^- = \sum_{i=1}^n d_i^+$ edges and D_n^- will be the random variable obtained by picking an element of $\mathbf{d}_n^-$ uniformly at random similarly define D_n^+.

The *directed configuration model* with degree sequence $\vec{\mathbf{d}}_n$, denoted $\mathbb{DCM}(\vec{\mathbf{d}}_n)$, is formed by taking d_i^- "in-stubs" and d_i^+ "out-stubs" associated with the vertex i for $i = 1, \ldots, n$, choosing a perfect matching from the set of out-stubs to the set of in-stubs uniformly at random to form the edges of $\mathbb{DCM}(\vec{\mathbf{d}}_n)$ directing them from the out-stub to the in-stub and contract the stubs associated with vertex i for each i to form the vertices.

In the directed configuration model there are now two types of components whose sizes cannot simply be deduced from the configuration model for graphs: the strong components and the weak components.

Strong Components

The study of the strong component structure of the directed configuration model was initiated by Cooper and Frieze [13]. They found the threshold for the existence of a giant strongly component under certain conditions was the point $Q = 0$ where we redefine

$$Q = \frac{1}{m} \sum_{i=1}^{n} d_i^- (d_i^+ - 1)$$

Here we use Q to draw the analogue with the threshold in the configuration model for graphs. In particular, this is a very similar threshold to the undirected one.

The conditions under which this result have been improved over the years. Graf [32] showed one can take a larger maximum degree than was assumed by Cooper and Frieze and later, Cai and Perarnau [11] showed that we only need to assume all second moments are bounded to draw the conclusion of Cooper and Frieze.

In this thesis we give the first result about the size of the largest strongly connected component where $Q = o(1)$. We define a pair of parameters related to the third moments of the degree sequence,

$$R^- := \frac{1}{m}\sum_{i=1}^{n} d_i^- d_i^+ (d_i^- - 1) \qquad\qquad R^+ := \frac{1}{m}\sum_{i=1}^{n} d_i^- d_i^+ (d_i^+ - 1)$$

We investigate the barely subcritical regime and prove the following

Theorem 1.3.3. Let $\mathbb{DCM}(\vec{\mathbf{d}}_n)$ be a random digraph from the directed configuration model with well behaved degree sequence $\vec{\mathbf{d}}_n$ and suppose that $nQ^3(R^-R^+)^{-1} \to -\infty$. Then a.a.s., there are no complex components or cycles of length $\omega(1/|Q|)$. Furthermore, the probability that the kth largest cycle, $\mathcal{C}_k$ has length at least $\alpha|Q|^{-1}$ is

$$\mathbb{P}\left(|\mathcal{C}_k| \geq \frac{\alpha}{|Q|}\right) = 1 - \sum_{i=0}^{k-1} \frac{\xi_\alpha^i}{i!} e^{-\xi_\alpha} + o(1),$$

where

$$\xi_\alpha = \int_\alpha^\infty \frac{e^{-x}}{x} dx.$$

This is an analogue of a similar theorem proved by Łuczak and Seierstad [55] for $D(n,p)$. We prove this theorem by splitting it into 4 pieces: cycles much longer than $1/|Q|$ which we show do not exist by an exploration process argument that there is no out-component so large; cycles slightly longer than $1/|Q|$ which we show are not present by a first moment argument; complex components are also shown not to exist by a first moment argument; finally to show the result on the length of the kth longest cycle we apply the Chen-Stein method for Poisson approximation. The details can be found in Chapter 4.

Weak components

Similarly to the strong components of the directed configuration model, the weak components have not received much attention. In this case it is perhaps due to the assumption that it effectively behaved like the undirected case (see, e.g., [64]). Kryven [49] observed that this assumption is wrong and predicted an alternative threshold for the appearance of the giant weak connected component supported with an analytical but non-rigorous approach based on generating functions for bounded directed degree distributions.

In this thesis we confirm that the location predicted by Kryven is correct (although we shall write the location differently). In particular, define $\mu_{i,j} = \mathbb{E}((D_n^-)_i(D_n^-)_j)$ (where $(x)_a = x(x-1)\dots(x-a+1)$) and $\lambda := \mu_{0,1} = \mu_{1,0}$. Also let

$$\rho = \frac{\mu_{1,1} + \sqrt{\mu_{2,0}\mu_{0,2}}}{\lambda}$$

Our main result is

Theorem 1.3.4. Let $\mathbb{DCM}(\vec{\mathbf{d}}_n)$ be a configuration model random digraph with degree sequence $\vec{\mathbf{d}}_n$. Then the point $\rho = 1$ is a threshold for the existence of a weakly connected component of linear order in $\mathbb{DCM}(\vec{\mathbf{d}}_n)$.

At first glance ρ may seem a mysterious parameter. The main idea of our proof is that ρ is the leading eigenvalue of the mean matrix of a 2-type Galton Watson process associated with the exploration of $\mathbb{DCM}(\vec{\mathbf{d}}_n)$. Thus, $\rho = 1$ is simply the point at which the associated process changes from subcritical to supercritical and thus should be where one would expect the location of the threshold to be. We give a full proof of this result in Chapter 5. This result comes from joint work with Guillem Perarnau. An extended abstract was published in [16].

CHAPTER 2

THE CRITICAL WINDOW OF BINOMIAL RANDOM DIGRAPHS

2.1 Introduction

Consider the random digraph model $D(n,p)$ where each of the $n(n-1)$ possible edges is included with probability p independently of all others. This is analogous to the Erdős-Renyi random graph $G(n,p)$ in which each edge is again present with probability p independently of all others. McDiarmid [57] showed that due to the similarity of the two models, it is often possible to couple $G(n,p)$ and $D(n,p)$ to compare the probabilities of certain properties.

In the random graph $G(n,p)$ the component structure is well understood. In their seminal paper [24], Erdős and Rényi proved that for $p = c/n$ the largest component of $G(n,p)$ has size $O(\log(n))$ if $c < 1$, is of order $\Theta(n^{2/3})$ if $c = 1$, and has linear size when $c > 1$. This threshold behaviour is known as the double jump. If we zoom in further around the critical point, $p = 1/n$ and consider $p = (1+\varepsilon(n))/n$ such that $\varepsilon(n) \to 0$ and $|\varepsilon(n)|^3 n \to \infty$, Bollobás [7] proved the following theorem for $|\varepsilon| > (2\log(n))^{1/2} n^{-1/3}$,which was extended to the whole range described above by Łuczak [52].

Theorem 2.1.1 ([7, 52])**.** Let $np = 1+\varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$ but $n|\varepsilon|^3 \to \infty$, and $k_0 = 2\varepsilon^{-2}\log(n|\varepsilon|^3)$.

i) If $n\varepsilon^3 \to -\infty$ then a.a.s. $G(n,p)$ contains no component of size greater than k_0.

ii) If $n\varepsilon^3 \to \infty$ then a.a.s. $G(n,p)$ contains a unique component of size greater than k_0. This component has size $2\varepsilon n(1+o(1))$.

Within the critical window itself i.e. $p = n^{-1} + \lambda n^{-4/3}$ with $\lambda \in \mathbb{R}$, the size of the largest component $\mathcal{C}_1$ is not tightly concentrated as it is for larger p. Instead, there exists a random variable $X_1 = X_1(\lambda)$ such that $|\mathcal{C}_1|n^{-2/3} \to X_1$ as $n \to \infty$. Much is known about the distribution of X_1, in fact the vector $\mathbf{X} = (X_1, \ldots, X_k)$ of normalised sizes of the largest k components i.e. $X_i = |\mathcal{C}_i|n^{-2/3}$ converges to the vector of longest excursion lengths of an inhomogeneous reflected Brownian motion by a result of Aldous [2]. In a more quantitative setting where one is more interested about behaviour for somewhat small n, Nachmias and Peres [63] proved the following (similar results may be found in [69, 73]).

Theorem 2.1.2 ([63])**.** Suppose $0 < \delta < 1/10$, $A > 8$ and n is sufficiently large with respect to A, δ. Then if $\mathcal{C}_1$ is the largest component of $G(n, 1/n)$, we have

i) $\mathbb{P}(|\mathcal{C}_1| < \lfloor \delta n^{2/3} \rfloor) \leq 15\delta^{3/5}$

ii) $\mathbb{P}(|\mathcal{C}_1| > An^{2/3}) \leq \frac{4}{A} e^{-\frac{A^2(A-4)}{32}}$

Note we have only stated the version of their theorem with $p = n^{-1}$ for clarity but it holds for the whole critical window. Of course, there are a vast number of other interesting properties of $\mathcal{C}_1$, see [1, 42, 54] for a number of examples.

In the setting of $D(n,p)$, one finds that analogues of many of the above theorems still hold. When working with digraphs, we are interested in the strongly connected components which we will often call the components. Note that the weak component structure of $D(n,p)$ is precisely the component structure of $G(n, 2p - p^2)$. For $p = c/n$, Karp [47] and Łuczak [53] independently showed that for $c < 1$ all components are of size $O(1)$ and when $c > 1$ there is a unique complex component of linear order and every other component is of size $O(1)$ (a component is complex if it has more edges than vertices).

The range $p = (1+\varepsilon)/n$ was studied by Łuczak and Seierstad [55] who were able to show the following result which can be viewed as a version of Theorem 2.1.1 for $D(n,p)$,

Theorem 2.1.3 ([55]). Let $np = 1+\varepsilon$, such that $\varepsilon = \varepsilon(n) \to 0$.

i) If $n\varepsilon^3 \to -\infty$ then a.a.s. every component of $D(n,p)$ is an isolated vertex or a cycle of length $O(1/|\varepsilon|)$.

ii) If $n\varepsilon^3 \to \infty$ then a.a.s. $D(n,p)$ contains a unique complex component of size $4\varepsilon^2 n(1+o(1))$ and every other component is an isolated vertex or a cycle of length $O(1/\varepsilon)$.

As a corollary Łuczak and Seierstad obtain a number of weaker results inside the critical window regarding complex components. They showed that there are $O_p(1)$ complex components containing $O_p(n^{1/3})$ vertices combined and that each has spread $\Omega_p(n^{1/3})$ (the *spread* of a complex digraph is the length of its shortest induced path).

Our main result is to give bounds on the tail probabilities of $|\mathcal{C}_1|$ resembling those of Nachmias and Peres [63] for $G(n,p)$.

Theorem 2.1.4 (Lower Bound). Let $0 < \delta < 1/800$, $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}$. Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then if n is sufficiently large with respect to δ, λ,

$$\mathbb{P}(|\mathcal{C}_1| < \delta n^{1/3}) \leq 2e\delta^{1/4}, \tag{2.1}$$

provided that $\delta \leq \frac{(\log 2)^2}{4|\lambda|^2}$.

Note that the constants in the above theorem have been chosen for simplicity and it is possible to give an expression for (2.1) depending on both λ and δ which imposes no restriction on their relation to one another.

Theorem 2.1.5 (Upper Bound). There exist constants, $\zeta, \eta > 0$ such that for any $A > 0, \lambda \in \mathbb{R}$, let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then

provided n is sufficiently large with respect to A, λ,

$$\mathbb{P}(|\mathcal{C}_1| > An^{1/3}) \leq \zeta e^{-\eta A^{3/2} + \lambda^+ A}$$

Where $\lambda^+ = \max(\lambda, 0)$.

A simple corollary of these bounds is that the largest component has size $\Theta(n^{1/3})$. This follows by taking $\delta = o(1)$ in Theorem 2.1.4 and $A = \omega(1)$ in Theorem 2.1.5.

Corollary 2.1.6. Let $\mathcal{C}_1$ be the largest component of $D(n,p)$ for $p = n^{-1} + \lambda n^{-4/3}$. Then, $|\mathcal{C}_1| = \Theta_p(n^{1/3})$.

Recently some related results have been obtained by Goldschmidt and Stephenson [31] who showed a scaling limit for the sizes of all the strong components in the critical random digraph. This is analogous to the result of Aldous [2] in $G(n,p)$ and allows one to deduce that there is a limiting distribution for the random variable $n^{-1/3}|\mathcal{C}_1|$. This distribution can be described in terms of the total edge length of a directed multigraph related to the Brownian continuum random tree. In principle versions of Theorems 2.1.4 and 2.1.5 can be deduced from this result although the limit distribution is a little difficult to work with. This author was only able to deduce that $n^{-1/3}|\mathcal{C}_1|$ is tight using the results of Goldschmidt and Stephenson whereas Theorems 2.1.4 and 2.1.5 give us much more explicit information about the tails of the random variable $n^{-1/3}|\mathcal{C}_1|$.

It should be noted that, in contrast to the undirected case, checking whether a set of W of vertices constitutes a strongly connected component of a digraph D requires much more than checking only those edges with at least one end in W. In particular, in order for W to be a strongly connected component, it must be strongly connected and there must be no directed path starting and ending in W which contains vertices that are not in W. This precludes us from using a number of methods which have often been used to study $G(n,p)$. We therefore develop novel methods for counting the number of strongly connected components of $D(n,p)$ based upon branching process arguments.

The remainder of this chapter is organised as follows. In Section 2.2 we give a pair of bounds on the number of strongly connected digraphs which have a given excess and number of vertices. Sections 2.3 and 2.4 contain the proofs of Theorems 2.1.4 and 2.1.5 respectively in the case that $p = n^{-1}$. The proof of Theorem 2.1.4 in Section 2.3 is a relatively straightforward application of Janson's inequality. The proof of Theorem 2.1.5 in Section 2.4 is much more involved. We use an exploration process to approximate the probability that a given subdigraph of $D(n,p)$ is also a component. Using this we approximate the expected number of strongly connected components of size at least $An^{1/3}$ and apply Markov's inequality. The adaptations required to handle the critical window $p = n^{-1} + \lambda n^{-4/3}$ are presented in Section 2.5. We conclude the chapeter in Section 2.6 with some open questions and final remarks.

2.2 Enumeration of Digraphs by size and excess

For both the upper and lower bounds on the size of the largest component, we need good bounds on the number strongly connected digraphs with a given excess and number of vertices. Where the *excess* of a strongly connected digraph with v vertices and e edges is $e - v$. Let $Y(m,k)$ be the number of strongly connected digraphs with m vertices and excess k. The study of $Y(m,k)$ was imitated by Wright [78] who obtained recurrences for the exact value of $Y(m,k)$. However, these recurrences swiftly become intractable as k grows. This has since been extended to asymptotic formulae when $k = \omega(1)$ and $O(m\log(m))$ [66, 71]. Note that when $k = m\log(m) + \omega(m)$, the fact $Y(m,k) \sim \binom{m(m-1)}{m+k}$ is a simple corollary of a result of Palásti [65]. In this section we give an universal bound on $Y(m,k)$ (Lemma 2.2.1) as well as a stronger bound for small excess (Lemma 2.2.3).

Lemma 2.2.1. *For every $m, k \geq 1$,*

$$Y(m,k) \leq \frac{(m+k)^k m^{2k}(m-1)!}{k!}$$

Proof. We will prove this by considering ear decompositions of the strongly connected digraphs in question. An *ear* is a non-trivial directed path in which the endpoints may coincide (i.e. it may be a cycle with a marked start/end vertex). The internal vertices of an ear are those that are not endpoints. An *ear decomposition* of a digraph D is a sequence, $E_0, E_1, \ldots, E_k$ of ears such that:

- E_0 is a cycle
- The endpoints of E_i belong to $\bigcup_{j=0}^{i-1} E_j$
- The internal vertices of E_i are disjoint from $\bigcup_{j=0}^{i-1} E_j$
- $\bigcup_{i=0}^{k} E_i = D$

We make use of the following fact.

Fact 2.2.2. A digraph D has an ear decomposition with $k+1$ ears if and only if D is strongly connected with excess k.

Thus we count strongly connected digraphs by a double counting of the number of possible ear decompositions. We produce an ear decomposition with m vertices and $k+1$ ears as follows. First, pick an ordering π of the vertices. Then insert k bars between the vertices such that the earliest the first bar may appear is after the second vertex in the order; multiple bars may be inserted between a pair of consecutive vertices. Finally, for each $i \in [k]$, we choose an ordered pair of vertices (u_i, v_i) which appear in the ordering before the ith bar.

This corresponds to a unique ear decomposition. The vertices in π before the first bar are E_0 with its endpoint being the first vertex. The internal vertices of E_i are the vertices of π between the ith and $i+1$st bar. Furthermore, E_i has endpoints u_i and v_i and is directed from u_i to v_i. The orientation of every other edge follows the order π.

Hence, there are at most

$$\binom{m+k-2}{k} m^{2k} m! \leq \frac{(m+k)^k m^{2k} m!}{k!}$$

ear decompositions. Note that each vertex of a strongly connected digraph is contained in a cycle. Therefore each vertex could be the endpoint of E_0 and hence at least m ear decompositions correspond to each strongly connected digraph. Hence the number of strongly connected digraphs of excess k may be bounded by

$$Y(m,k) \leq \frac{(m+k)^k m^{2k} m!}{k! m} = \frac{(m+k)^k m^{2k}(m-1)!}{k!},$$

as claimed. □

Lemma 2.2.3. *There exists $C > 0$ such that for $1 \leq k \leq \sqrt{m}/3$ and m sufficiently large we have,*

$$Y(m,k) \leq C \frac{m! m^{3k-1}}{(2k-1)!}. \tag{2.2}$$

The proof of the above lemma follows similar lines to the proof of Theorem 1.1 in [66] to obtain a bound of a similar order. We then prove that this bound implies the above which is much easier to work with.

First we introduce some definitions and notation from [66]. A random variable X has the zero-truncated Poisson distribution with parameter $\lambda > 0$ denoted $X \sim TP(\lambda)$ if it has probability mass function

$$\mathbb{P}(X = i) = \begin{cases} \frac{\lambda^i}{i!(e^\lambda - 1)} & \text{if } i \geq 1, \\ 0 & \text{if } i < 1. \end{cases}$$

Let $\mathcal{D}$ be the collection of all degree sequences $\mathbf{d} = (d_1^+, \ldots, d_m^+, d_1^-, \ldots, d_m^-)$ such that $d_i^+, d_i^- \geq 1$ for each $1 \leq i \leq m$ and furthermore,

$$\sum_{i=1}^m d_i^+ = \sum_{i=1}^m d_i^- = m.$$

A *preheart* is a digraph with minimum semi-degree at least 1 and no cycle components.

The *heart* of a preheart D is the multidigraph $H(D)$ formed by suppressing all vertices of D which have in and out degree precisely 1.

We define the *preheart configuration model*, a two stage variant of the configuration model for digraphs which always produces a preheart, as follows. For $\mathbf{d} \in \mathcal{D}$, define

$$T = T(\mathbf{d}) = \{i \in [m] : d_i^+ + d_i^- \geq 3\}.$$

First we apply the configuration model to T to produce a heart H. That is, assign each vertex $i \in T$ d_i^+ out-stubs and d_i^- in-stubs and pick a uniformly random perfect matching between in- and out-stubs. Next, given a heart configuration H, we construct a preheart configuration Q by assigning $[m] \setminus T$ to $E(H)$ such that the vertices assigned to each arc of H are given a linear order. Denote this assignment including the orderings by q. Then the preheart configuration model, $\mathcal{Q}(\mathbf{d})$ is the probability space of random preheart configurations formed by choosing H and q uniformly at random. Note that each $Q \in \mathcal{Q}(\mathbf{d})$ corresponds to a (multi)digraph with m vertices $m + k$ edges and degree sequence $\mathbf{d}$.

As in the configuration model, each simple digraph with degree sequence $\mathbf{d}$ is produced in precisely $\prod_{i=1}^{m} d_i^+! d_i^-!$ ways. So if we restrict to simple preheart configurations, the digraphs we generate in this way are uniformly distributed. Where in this case, simple means that there are no multiple edges or loops (however cycles of length 2 are allowed). We now count the number of preheart configurations. Let $m' = m'(\mathbf{d}) = |T(\mathbf{d})|$ be the number of vertices of the heart. Then, we have the following

Lemma 2.2.4. *Let $\mathbf{d} \in \mathcal{D}$, then there are*

$$\frac{m'(\mathbf{d}) + k}{m + k}(m + k)!$$

preheart configurations.

Proof. We first generate the heart, and as we are simply working with the configuration

model for this part of the model, there are $(m'+k)!$ heart configurations. The assignment of vertices in $[m] \setminus T$ to the arcs of the heart H may be done one vertex at a time by subdividing any already present edge and maintaining orientation. In this way when we add the ith vertex in this stage, there are $m'+k+i-1$ choices for the edge we subdivide. We must add $m-m'$ edges in this stage and so there are

$$\prod_{i=1}^{m-m'} m'+k+i-1 = \frac{(m+k-1)!}{(m'+k-1)!}$$

unique ways to create a preheart configuration from any given heart. Multiplying the number of heart configurations by the number of ways to create a preheart configuration from a given heart yields the desired result. □

The next stage is to pick the degree sequence, $\mathbf{d} \in \mathcal{D}$ at random. We do this by choosing the degrees to be independent and identically distributed zero-truncated Poisson random variables with mean $\lambda > 0$. That is, $d_i^+ \sim TP(\lambda)$ and $d_i^- \sim TP(\lambda)$ such that the family $\{d_i^+, d_i^- : i \in [m]\}$ is independent. Note that this may not give a degree sequence at all, or it may be the degree sequence of a digraph with the wrong number of edges. Thus we define the event $\Sigma(\lambda)$ to be the event that

$$\sum_{i=1}^{m} d_i^+ = \sum_{i=1}^{m} d_i^- = m+k.$$

We shall now prove the following bound,

Lemma 2.2.5. *For any $\lambda > 0$ we have*

$$Y(m,k) \leq \frac{3k(m+k-1)!(e^\lambda-1)^{2m}}{\lambda^{2(m+k)}} \mathbb{P}(\Sigma(\lambda)). \tag{2.3}$$

Proof. Let $\mathbf{D}$ be the random degree sequence generated as above and $\mathbf{d} \in \mathcal{D}$, then

$$\mathbb{P}(\mathbf{D}=\mathbf{d}) = \prod_{i=1}^{m} \frac{\lambda^{d_i^+}}{d_i^+!(e^\lambda-1)} \frac{\lambda^{d_i^-}}{d_i^-!(e^\lambda-1)} = \frac{\lambda^{2(m+k)}}{(e^\lambda-1)^{2m}} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!}. \tag{2.4}$$

By definition of $\Sigma(\lambda)$, we have

$$\sum_{\mathbf{d}\in\mathcal{D}} \mathbb{P}(\mathbf{D}=\mathbf{d}) = \mathbb{P}(\Sigma(\lambda)),$$

as all of the above events are disjoint. Thus, we may rearrange (2.4) to deduce that

$$\sum_{\mathbf{d}\in\mathcal{D}} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!} = \frac{(e^\lambda-1)^{2m}}{\lambda^{2(m+k)}}\mathbb{P}(\Sigma(\lambda)). \tag{2.5}$$

Lemma 2.2.4 tells us that for a given degree sequence $\mathbf{d}$, there are

$$\frac{m'(\mathbf{d})+k}{m+k}(m+k)!$$

preheart configurations. As each simple digraph with degree sequence $\mathbf{d}$ comes from precisely $\prod_{i=1}^{m} d_i^+!d_i^-!$ configurations, and $m'(\mathbf{d}) \leq 2k$ as otherwise the excess would be larger than k, we can deduce that the total number of prehearts with m vertices and excess k is

$$\sum_{\mathbf{d}\in\mathcal{D}} \frac{m'(\mathbf{d})+k}{m+k}(m+k)! \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!} \leq \sum_{\mathbf{d}\in\mathcal{D}} (m+k)! \frac{3k}{m+k} \prod_{i=1}^{m} \frac{1}{d_i^+!d_i^-!}. \tag{2.6}$$

Note that any strongly connected digraph is a preheart and so (2.6) is also an upper bound for $Y(m,k)$. Finally, combining (2.5) and (2.6) yields the desired inequality. □

It remains to prove that (2.3) can be bounded from above by (2.2). To this end, we prove the following upper bound on $\mathbb{P}(\Sigma(\lambda))$.

Lemma 2.2.6. *For $\lambda < 1$,*

$$\mathbb{P}(\Sigma(\lambda)) \leq \frac{147}{\lambda m}.$$

For the proof of this lemma, we will use the Berry-Esseen inequality for normal approximation (see for example [75, Section XX.2].)

Lemma 2.2.7. *Suppose $X_1, X_2, \ldots, X_n$ is a sequence of independent random variables from a common distribution with zero mean, unit variance and third absolute moment $\mathbb{E}|X|^3 = \gamma < \infty$. Let $S_n = X_1 + X_2 + \ldots + X_n$ and let G_n be the cumulative distribution function of $S_n/\sqrt{n}$. Then for each n we have*

$$\sup_{t \in \mathbb{R}} |G_n(t) - \Phi(t)| \leq \frac{\gamma}{2\sqrt{n}}, \tag{2.7}$$

where Φ is the cumulative distribution function of the standard Gaussian.

Here, the explicit constant 1/2 in equation (2.7) was obtained by Tyurin [74].

Proof of Lemma 2.2.6. The in-degrees of the random degree sequence are chosen independently from a truncated poisson distribution with parameter λ. Thus, we want to apply Lemma 2.2.7 to the sum $S_m = Y_1 + Y_2 + \ldots + Y_m$ where the Y_i are normalised truncated Poisson random variables. So all we must compute are the first three central moments of the truncated poisson distribution. Let $Y \sim TP(\lambda)$, one can easily compute that $\mathbb{E}(Y) = c_\lambda = \frac{\lambda e^\lambda}{e^\lambda - 1}$ and $\mathrm{Var}(Y) = \sigma_\lambda^2 = c_\lambda(1 + \lambda - c_\lambda)$. Note that for $\lambda < 1$ we have $1 < c_\lambda < 2$ and so as Y only takes integer values which are at least 1, $\mathbb{E}|Y - \mathbb{E}(Y)|^3 = \mathbb{E}(Y - c_\lambda)^3 + 2(c_\lambda - 1)^3 \mathbb{P}(Y = 1)$. Computing this yields

$$\mathbb{E}|Y - \mathbb{E}(Y)|^3 = \lambda + \frac{2\lambda^4 - 5\lambda^3 + 3\lambda^2 - \lambda}{e^\lambda - 1} + \frac{3(2\lambda^4 - 3\lambda^3 + \lambda^2)}{(e^\lambda - 1)^2} + \frac{2(3\lambda^4 - 2\lambda^3)}{(e^\lambda - 1)^3} + \frac{2\lambda^4}{(e^\lambda - 1)^4}$$

One can check that this is bounded above by 2λ for $\lambda < 1$.

The normalised version of Y is $X = (Y - c_\lambda)/\sigma_\lambda$. We have

$$\mathbb{E}|X|^3 = \mathbb{E}\left|\frac{Y - c_\lambda}{\sigma_\lambda}\right|^3 = \frac{1}{\sigma_\lambda^3}\mathbb{E}|Y - c_\lambda|^3 \leq \frac{2\lambda}{\sigma_\lambda^3} = \gamma.$$

For $\lambda < 1$ one can check $c_\lambda < 1 + 2\lambda/3$, which allows us to deduce that $\sigma_\lambda^2 > \lambda/3$ (also using $Y \geq 1$). Hence, $\mathbb{E}|X|^3 \leq 6\sqrt{3}\lambda^{-1/2}$. Substituting into Lemma 2.2.7 with G_m the

distribution of $S_m/\sqrt{m}$,

$$\sup_{t\in\mathbb{R}} |G_m(t) - \Phi(t)| \leq \frac{3\sqrt{3}}{\sqrt{\lambda m}}.$$

The probability that the sum of the in-degrees is $m+k$ is precisely

$$G_m\left(\frac{m+k-mc_\lambda}{\sigma_\lambda\sqrt{m}}\right) - G_m\left(\frac{m+k-1-mc_\lambda}{\sigma_\lambda\sqrt{m}}\right).$$

Following an application of the triangle inequality, we see that this probability is bounded above by

$$\frac{6\sqrt{3}}{\sqrt{\lambda m}} + \frac{1}{\sqrt{2\pi m}\sigma_\lambda} \leq \frac{7\sqrt{3}}{\sqrt{\lambda m}}.$$

As the event that the in-degrees sum to $m+k$ and the event that the out-degrees sum to $m+k$ are independent and identically distributed events, we may deduce the bound,

$$\mathbb{P}(\Sigma(\lambda)) \leq \frac{147}{\lambda m}.$$

□

Finally, we may prove Lemma 2.2.3.

Proof of Lemma 2.2.3. We choose $\lambda = 2k/m < 1$ by assumption, then $\mathbb{P}(\Sigma(\lambda)) \leq 147/2k$ by Lemma 2.2.5. Combining this with Lemma 2.2.6 yields

$$Y(m,k) \leq \frac{441(m+k-1)!}{2}\lambda^{-2k}\left(\frac{e^\lambda - 1}{\lambda}\right)^{2m} \leq \frac{441 m! m^{3k-1} e^{k^2/m}}{(2k)^{2k}}\left(\frac{e^\lambda - 1}{\lambda}\right)^{2m} \tag{2.8}$$

We use the inequality $e^x \leq 1 + x + x^2/2 + x^3/4$ which holds for all $0 \leq x \leq 1$ to bound $(e^\lambda - 1)/\lambda \leq 1 + \lambda/2 + \lambda^2/4$. Thus,

$$((e^\lambda - 1)/\lambda)^{2m} \leq (1 + \lambda/2 + \lambda^2/4)^{2m} \leq e^{m\lambda + m\lambda^2/2} = e^{2k+2k^2/m}.$$

Then, we can use Stirling's inequality, $e\sqrt{2k-1}(2k-1)^{2k-1}e^{-2k+1} \geq (2k-1)!$, so that

$$\frac{e^{2k}}{(2k)^{2k}} \leq \frac{e^{2k}}{(2k-1)^{2k-1/2}} \leq \frac{e^2}{(2k-1)!},$$

allowing us to rewrite the bound on $Y(m,k)$ as

$$Y(m,k) \leq \frac{441e^3}{2}\frac{m!m^{3k-1}}{(2k-1)!},$$

where we used $e^{k^2/m} \leq e^{1/3}$. Thus proving the lemma with $C = 441e^3/2$. □

2.3 Proof of Theorem 2.1.4

In this section we prove a lower bound on component sizes in $D(n,p)$. We give the proof for $p = 1/n$ for simplicity. The proof when $p = n^{-1} + \lambda n^{-4/3}$ is very similar, with more care taken in the approximation of terms involving $(np)^m$. See Section 2.5 for more details.

Theorem 2.3.1. Let $0 < \delta < 1/800$, then the probability that $D(n, 1/n)$ has no component of size at least $\delta n^{1/3}$ is at most $2\delta^{1/2}$.

To prove this we will bound from above the probability that there is no cycle of length between $\delta n^{1/3}$ and $\delta^{1/2}n^{1/3}$. Let X be the random variable counting the number of cycles in $D(n, 1/n)$ of length between $\delta n^{1/3}$ and $\delta^{1/2}n^{1/3}$. Note that we may decompose X as a sum of dependent Bernoulli random variables, and thus we may apply Janson's Inequality in the following form (see [44, Theorem 2.18 (i)]).

Theorem 2.3.2. Let S be a set and $S_p \subseteq S$ chosen by including each element of S in S_p independently with probability p. Suppose that $\mathcal{S}$ is a family of subsets of S and for $A \in \mathcal{S}$, we define I_A to be the event $\{A \subseteq S_p\}$. Let $\mu = \mathbb{E}(X)$ and

$$\Delta = \frac{1}{2}\sum\sum_{A \neq B, A \cap B \neq \emptyset} \mathbb{E}(I_A I_B)$$

Then,

$$\mathbb{P}(X = 0) \leq e^{-\mu+\Delta}$$

To apply Theorem 2.3.2, we define S to be the set of edges of the complete digraph on n vertices. Let $A \in \mathcal{S}$ if and only if $A \subseteq S$ is the set of edges of a cycle of length between $\delta n^{1/3}$ and $\delta^{1/2} n^{1/3}$. Define $X(m)$ to be the number cycles in $D(n, 1/n)$ of length m. We start by approximating the first moment of X.

Lemma 2.3.3. $\mathbb{E}(X) \geq \log(1/\delta)/2$

Proof. Let $a = \delta n^{1/3}$ and $b = \delta^{1/2} n^{1/3}$. Then, we can write X as

$$X = \sum_{m=a}^{b} X(m)$$

Note that

$$\mathbb{E}(X(m)) = \binom{n}{m} \frac{m!}{m} p^m \geq \frac{1}{m} \tag{2.9}$$

So, we may bound the expectation of X as follows

$$\mathbb{E}(X) = \sum_{m=a}^{b} \mathbb{E}(X(m)) \geq \sum_{m=a}^{b} \frac{1}{m} \geq \int_a^b \frac{dx}{x} = \frac{\log(1/\delta)}{2}$$

□

Let $Z(m, k)$ be the random variable counting the number of strongly connected graphs with m vertices and excess k in $D(n, 1/n)$. Directly computing Δ is rather complicated so we will instead compute an upper bound on Δ that is a linear combination of the first moments of the random variables $Z(m, k)$ for $m \geq a$ and $k \geq 1$. To move from the computation of Δ to the first moments of $Z(m, k)$ we use the following lemma,

Lemma 2.3.4. *Each strongly connected digraph D with excess k may be formed in at most 27^k ways as the union of a pair of directed cycles C_1 and C_2.*

Proof. Consider the heart $H(D)$ of D. Recall that $H(D)$ is the (multi)-digraph formed by suppressing the degree 2 vertices of D and retaining orientations. As D has excess

k, $H(D)$ has at most $2k$ vertices. Furthermore, the excess of $H(D)$ is the same as the excess of D as we only remove vertices of degree 2. Thus $H(D)$ has at most $3k$ edges.

Then, each edge of $H(D)$ must be a subdigraph of either C_1, C_2 or both. So there are $3^{3k} = 27^k$ choices for the pair C_1, C_2 as claimed. $\square$

We are now in a position to give a bound on Δ.

Lemma 2.3.5. $\Delta \leq \log(2)$ *for any* $\delta \in (0, 1/800]$

Proof. Let

$$\Gamma(k) := \{E(C) | C \subseteq \overrightarrow{K_n}, C \cong \overrightarrow{C_k}\},$$

where $\overrightarrow{K_n}$ is the complete digraph on $[n]$ and $\overrightarrow{C_k}$ is the directed cycle of length k. For $\alpha \in \Gamma(k)$ let I_α be the indicator function of the event that all edges of α are present in a given realisation of $D(n, 1/n)$. Also, define

$$\Gamma = \bigcup_{k=a}^{b} \Gamma(k).$$

Then, by definition,

$$\Delta = \frac{1}{2} \sum\sum_{\alpha \neq \beta, \alpha \cap \beta \neq \emptyset} \mathbb{E}(I_\alpha I_\beta)$$

Let $\Gamma_\alpha^{m,k}(t)$ be the set of $\beta \in \Gamma(t)$ such that $\alpha \cup \beta$ is a collection of $m + k$ edges spanning m vertices. Then,

$$\begin{aligned}
2\Delta &= \sum_{s=a}^{b} \sum_{t=a}^{b} \sum_{\alpha \in \Gamma(s)} \sum_{m=s}^{\infty} \sum_{k=1}^{\infty} \sum_{\beta \in \Gamma_\alpha^{m,k}(t)} p^{m+k} \\
&\leq \sum_{m=a}^{2b} \sum_{k=1}^{\infty} \sum_{s=a}^{m} \sum_{t=a}^{m} \sum_{\alpha \in \Gamma(s)} \sum_{\beta \in \Gamma_\alpha^{m,k}(t)} p^{m+k} \\
&\leq \sum_{m=a}^{2b} \sum_{k=1}^{\infty} 27^k \mathbb{E}(Z(m,k)), \qquad (2.10)
\end{aligned}$$

where the last inequality follows from Lemma 2.3.4. Note that

$$\mathbb{E}(Z(m,k)) = \binom{n}{m} p^{m+k} Y(m,k)$$

by definition. We will use the following two bounds on $Y(m,k)$ which follow immediately from Lemma 2.2.1.

- If $k \leq m$, then $Y(m,k) \leq \frac{2^k m^{3k} m!}{k! m}$
- If $k > m$, then $Y(m,k) \leq \frac{(2e)^k m^{2k} m!}{m}$

This allows us to split the sum in (2.10) based upon whether $k \leq m$ or $k > m$ to obtain

$$\begin{aligned}
2\Delta &\leq \sum_{m=a}^{2b} \sum_{k=1}^{m} 27^k \binom{n}{m} \frac{2^k m^{3k} m!}{k! m} p^{m+k} + \sum_{m=a}^{2b} \sum_{k=m+1}^{\infty} 27^k \binom{n}{m} \frac{(2e)^k m^{2k} m!}{m} p^{m+k} \\
&\leq \sum_{m=a}^{2b} \frac{1}{m} \sum_{k=1}^{\infty} \frac{(54pm^3)^k}{k!} + \sum_{m=a}^{2b} \frac{1}{m} \sum_{k=m+1}^{\infty} (54em^2p)^k \\
&\leq \frac{\log(4/\delta)}{2} (e^{432\delta^{3/2}} - 1 + 23328e^2\delta^2) \qquad (2.11)
\end{aligned}$$

Where the $23328e^2\delta^2$ term comes from noting $k \geq 2$ in the range $k \geq m+1$ and that for $x \leq 1/2$

$$\sum_{k=2}^{\infty} x^k \leq 2x^2$$

As (2.11) is increasing in δ, we simply need to check that the Lemma holds for $\delta = 1/800$ which may be done numerically.

□

Finally, to prove Theorem 2.3.1 we substitute the values obtained for μ and Δ in Lemmas 2.3.3 and 2.3.5 respectively into Theorem 2.3.2. That is,

$$\mathbb{P}(X = 0) \leq e^{-\mu+\Delta} \leq e^{-\log(1/\delta)/2+\log(2)} = 2\delta^{1/2}$$

So the probability there is no directed cycle of length at least $\delta n^{1/3}$ is at most $2\delta^{1/2}$ and,

as cycles are strongly connected, this is also an upper bound on the probability there is no strongly connected component of size at least $\delta n^{1/3}$.

2.4 Proof of Theorem 2.1.5

In this section we prove an upper bound on the component sizes in $D(n,p)$. Again, we only consider the case when $p = 1/n$ to simplify notation and calculations. The reader is referred to Section 2.5 for a sketch of the adaptations to extend the result to the full critical window. The following is a restatement of Theorem 2.1.5 for $p = 1/n$.

Theorem 2.4.1. There exist constants $\zeta, \eta > 0$ such that for any $A > 0$ if n is sufficiently large with respect to A, then the probability that $D(n, 1/n)$ contains any component of size at least $An^{1/3}$ is at most $\zeta e^{-\eta A^{3/2}}$.

We will use the first moment method to prove this theorem and calculate the expected number of large strongly connected components in $D(n, 1/n)$. Note that it is important to count components and not strongly connected subgraphs as the expected number of strongly connected subgraphs in $D(n, 1/n)$ blows up as $n \to \infty$. Thus for each strongly connected subgraph, we will use an exploration process to determine whether or not it is a component.

The exploration process we use was initially developed by Martin-Löf [56] and Karp [47]. During this process, vertices will be in one of three classes: *active*, *explored* or *unexplored*. At time $t \in \mathbb{N}$, we let X_t be the number of active vertices, A_t the set of active vertices, E_t the set of explored vertices and U_t the set of unexplored vertices.

We will start from a set A_0 of vertices of size X_0 and fix an ordering of the vertices, starting with A_0. For step $t \geq 1$, if $X_{t-1} > 0$ let w_t be the first active vertex. Otherwise, let w_t be the first unexplored vertex. Define η_t to be the number of unexplored out-neighbours of w_t in $D(n, 1/n)$. Change the class of each of these vertices to active and set w_t to explored. This means that $|E_t| = t$ and furthermore, $|U_t| = n - X_t - t$. Let

$N_t = n - X_t - t - \mathbb{1}(X_t = 0)$ be the number of potential unexplored out-neighbours of w_{t+1} i.e. the number of unexplored vertices which are not w_{t+1}. Then, given the history of the process, η_t is distributed as a binomial random variable with parameters N_{t-1} and $1/n$. Furthermore, the following recurrence relation holds.

$$X_t = \begin{cases} X_{t-1} + \eta_t - 1 & \text{if } X_{t-1} > 0, \\ \eta_t & \text{otherwise} \end{cases} \tag{2.12}$$

Let $\tau_1 = \min\{t \geq 1 : X_t = 0\}$. Note that this is a stopping time and at time τ_1 the set E_{τ_1} of explored vertices is precisely the out-component of A_0. If A_0 spans a strongly connected subdigraph D_0 of $D(n, 1/n)$, then D_0 is a strongly connected component if and only if there are no edges from $E_{\tau_1} \setminus A_0$ to A_0. The key idea will be to show that if X_0 is sufficiently large, then it is very unlikely for τ_1 to be small, and consequently it is also very unlikely that there are no edges from $E_{\tau_1} \setminus A_0$ to A_0. This is encapsulated in the following lemma.

Lemma 2.4.2. *Let X_t be the exploration process defined above with starting set of vertices A_0 of size $X_0 = m$. Suppose $0 < c < \sqrt{2}$ is a fixed constant. Then,*

$$\mathbb{P}(\tau_1 < cm^{1/2}n^{1/2}) \leq 2e^{-\frac{(2-c^2)^2}{8c}m^{3/2}n^{-1/2}+O(m^2n^{-1})}.$$

Proof. Define $\xi = cm^{1/2}n^{1/2}$ and consider the auxiliary process, X'_t which we define recursively by

$$X'_0 = m,$$
$$X'_t = X'_{t-1} - 1 + W_t \text{ for } t \geq 1,$$

where $W_t \sim \text{Bin}(n - t - 10m, p)$. Let τ_2 be the stopping time,

$$\tau_2 = \inf\{t : X_t > 10m\}$$

We may couple the processes (X_t, X'_t) such that X'_t is stochastically dominated by X_t for $t < \tau_2$. The coupling may be explicitly defined by setting $\eta_t = W_t + W'_t$ with $W'_t \sim \mathrm{Bin}(10m - X_{t-1}, p)$. Define another stopping time, $\tau'_1 = \min\{t \geq 1 : X'_t = 0\}$. Consider the following events

$$\mathcal{E}_1 = \{\tau_1 < cm^{1/2}n^{1/2}\}$$
$$\mathcal{E}_2 = \{\tau'_1 < cm^{1/2}n^{1/2}\}$$
$$\mathcal{E}_3 = \{\tau_2 < cm^{1/2}n^{1/2}\}$$

And note that $\mathbb{P}(\mathcal{E}_1) \leq \mathbb{P}(\mathcal{E}_2) + \mathbb{P}(\mathcal{E}_3)$ by our choice of coupling and a union bound (as the coupling guarantees $\mathcal{E}_1 \subseteq \mathcal{E}_2 \cup \mathcal{E}_3$). Thus we only need to bound the probabilities of the simpler events $\mathcal{E}_2$ and $\mathcal{E}_3$. We begin by considering $\mathcal{E}_3$. To bound its probability we consider the upper bound process M_t defined by

$$M_0 = m,$$
$$M_t = M_{t-1} - 1 + B_t \text{ for } t \geq 1,$$

where $B_t \sim \mathrm{Bin}(n, 1/n)$. It is straightforward to couple (X_t, M_t) such that M_t stochastically dominates X_t. Furthermore, M_t is a martingale. Hence, $\mathbb{P}(\mathcal{E}_3) \leq \mathbb{P}(\tau'_2 < cm^{1/2}n^{1/2})$ where τ'_2 is the stopping time, $\tau'_2 = \min\{t : M_t > 10m\}$. To bound the probability of $\mathcal{E}_2$ consider the process Y_t defined as $Y_t = m - X'_t$. One can check that Y_t is a submartingale.

As $x \mapsto e^{\alpha x}$ is a convex non-decreasing function for any $\alpha > 0$, we may apply Jensen's inequality to deduce that $Z_t^- = e^{\alpha Y_t}$ and $Z_t^+ = e^{\alpha M_t}$ are submartingales. Also, $Z_t^-, Z_t^+ > 0$ for any $i \in \mathbb{N}$. Starting with Z_t^-, we may apply Doob's maximal inequality [33, Section 12.6] and deduce that

$$\mathbb{P}\left(\min_{0 \leq t \leq \xi} X'_i \leq 0\right) = \mathbb{P}\left(\max_{0 \leq t \leq \xi} Z_t^- \geq e^{\alpha m}\right) \leq \frac{\mathbb{E}(Z_\xi^-)}{e^{\alpha m}} \tag{2.13}$$

We may rewrite this by noting that

$$Y_t = m - X_t' = t - \sum_{i=1}^{t} W_i = t - R_t$$

where R_t is binomially distributed and in particular $R_\xi \sim \mathrm{Bin}(l\xi, p)$ for

$$l\xi = cm^{1/2}n^{3/2} - \frac{c^2mn}{2} - 10cm^{3/2}n^{1/2} + \frac{cm^{1/2}n^{1/2}}{2}$$

Also, we choose x such that $xl\xi = \xi - m$. Then (2.13) may be rewritten as $e^{-\alpha m}\mathbb{E}(Z_\xi^-) = e^{\alpha x l\xi}\mathbb{E}(e^{-\alpha R_\xi})$. The next stage is to rearrange this into a form which resembles the usual Chernoff bounds (for $x < p$). So, let

$$f(\alpha) = e^{\alpha x l\xi}\mathbb{E}(e^{-\alpha R_\xi}) = \left[e^{\alpha x}(pe^{-\alpha} + 1 - p)\right]^{l\xi}$$

Then, we choose α^* to minimise f. Solving $f'(\alpha) = 0$, we obtain the solution

$$e^{-\alpha^*} = \frac{x(1-p)}{p(1-x)}$$

Note $x < p$ so, $e^{-\alpha^*} < 1$ and $\alpha^* > 0$ as desired. Thus,

$$\begin{aligned} f(\alpha^*) = &= \left[\left(\frac{p(1-x)}{x(1-p)}\right)^x \left(x\frac{1-p}{1-x} + 1 - p\right)\right]^{mt} \\ &= \left[\left(x\frac{1-p}{1-x} + 1 - p\right)\left(\frac{p}{x}\right)^x \left(\frac{1-p}{1-x}\right)^x\right]^{mt} \\ &= \left[\left(\frac{p}{x}\right)^x \left(\frac{1-p}{1-x}\right)^{1-x}\right]^{mt} \end{aligned}$$

Which is the usual expression found in Chernoff bounds. As usual, we bound this by writing

$$f(\alpha^*) = e^{-g(x)l\xi}$$

and bound g, where

$$g(x) = x \log\left(\frac{x}{p}\right) + (1-x)\log\left(\frac{1-x}{1-p}\right)$$

Computing the Taylor expansion of g we find that $g(p) = g'(p) = 0$. So, if $g''(x) \geq \beta$ for all x between p and $p-h$, then $g(p-h) \geq \beta h^2/2$. Furthermore,

$$g''(x) = \frac{1}{x} + \frac{1}{1-x}$$

As $0 < x < p$, we have $g''(x) \geq 1/x \geq 1/p$. So, we deduce that $g(x) \geq \delta^2 p/2$ where $\delta = 1 - x/p$. All that remains is to compute δ. As defined earlier, we have $xl\xi = \xi - m$ which for convenience we will write as

$$xl\xi = \xi\left(1 - \frac{m^{1/2}}{cn^{1/2}}\right) \tag{2.14}$$

Also, as $p = n^{-1}$, and recalling the definition of $l\xi$ from earlier,

$$\begin{aligned} pl\xi &= cm^{1/2}n^{1/2} - \frac{c^2 m}{2} + O(m^{3/2}n^{-1/2}) \\ &= \xi\left(1 - \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1})\right) \end{aligned} \tag{2.15}$$

We divide (2.14) by (2.15) and as the Taylor expansion of $1/(1-w)$ is $\sum_{i\geq 0} w^i$,

$$\frac{x}{p} = \frac{1 - \frac{m^{1/2}}{cn^{1/2}}}{1 - \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1})} = 1 - \frac{m^{1/2}}{cn^{1/2}} + \frac{cm^{1/2}}{2n^{1/2}} + O(mn^{-1}) \tag{2.16}$$

From which we may deduce

$$\delta = \frac{(2-c^2)m^{1/2}}{2cn^{1/2}} + O(mn^{-1}) \tag{2.17}$$

So,

$$\mathbb{P}(\mathcal{E}_2) \leq e^{-\frac{\delta^2 p}{2} l\xi} = e^{-\frac{(2-c^2)^2}{8c} m^{3/2}n^{-1/2} + O(m^2 n^{-1})} \tag{2.18}$$

We may proceed similarly for Z_t^+, in particular we must still appeal to Doob's maximal inequality as we seek a bound over the entire process. In this case we end up with a $\mathrm{Bin}(n\xi, p)$ distribution and are looking at the upper tail rather than the lower. We find $pn\xi = \xi$ and

$$xn\xi = \xi + 9m = \xi\left(1 + \frac{9m^{1/2}}{cn^{1/2}}\right)$$

Thus,

$$\delta = \frac{x}{p} - 1 = \frac{9m^{1/2}}{cn^{1/2}}$$

Substituting into the analogous bound,

$$\mathbb{P}(\mathcal{E}_3) \leq e^{-\frac{\delta^2 p}{3} n\xi} \leq e^{-\frac{27m^{3/2}}{cn^{1/2}}} \tag{2.19}$$

Observe that $\mathbb{P}(\mathcal{E}_2) \geq \mathbb{P}(\mathcal{E}_3)e^{O(m^2n^{-1})}$ for $0 < c < \sqrt{2(1+3\sqrt{6})}$. Thus, in the range we are interested in, we may use $2\mathbb{P}(\mathcal{E}_2)$ as an upper bound for $\mathbb{P}(\mathcal{E}_2) + \mathbb{P}(\mathcal{E}_3)$ and this proves the lemma. □

We now compute the probability that any given strongly connected subgraph of $D(n, 1/n)$ is a component. To do so, we use the simple observation that a strongly connected subgraph is a component if it is not contained in a larger strongly connected subgraph.

Lemma 2.4.3. *There exist $\beta, \gamma > 0$ such that if H is any strongly connected subgraph of $D(n, 1/n)$ with m vertices. Then the probability that H is a strongly connected component of $D(n, 1/n)$ is at most $\beta e^{-(1+\gamma)m^{3/2}n^{-1/2} + O(m^2n^{-1})}$.*

Proof. We compute the probability that H is a component of $D(n, 1/n)$ by running the exploration process X_t starting from $A_0 = V(H)$. So, $X_0 = m$. Once the exploration process dies at time τ_1, any backward edge from $E_{\tau_1} \setminus A_0$ to A_0 gives a strongly connected subgraph of $D(n, 1/n)$ which contains H. Let Y_t be the random variable which counts the number of edges from $E_{\tau_1} \setminus A_0$ to A_0. Note that for $t \geq m$, $Y_t \sim \mathrm{Bin}(m(t-m), p)$. Furthermore, H is a strongly connected component of $D(n, 1/n)$ if and only if $Y_{\tau_1} = 0$.

Let $\varepsilon > 0$ and define the events $\mathcal{A}_i$ for $i = 1, \ldots, r$ (where $r \sim c/\varepsilon$ for some $c > 1$) to be

$$\mathcal{A}_i = \{(i-1)\varepsilon m^{1/2} n^{1/2} \leq \tau_1 < i\varepsilon m^{1/2} n^{1/2}\},$$
$$\mathcal{A}_{r+1} = \{r\varepsilon m^{1/2} n^{1/2} \leq \tau_1\}.$$

Clearly the family $\{\mathcal{A}_i : i = 1, \ldots, r+1\}$ forms a partition of the sample space. So, by the law of total probability,

$$\mathbb{P}(Y_{\tau_1} = 0) = \sum_{i=1}^{r+1} \mathbb{P}(Y_{\tau_1} = 0 | \mathcal{A}_i)\mathbb{P}(\mathcal{A}_i) \tag{2.20}$$

By applying Lemma 2.4.2 when $1 \leq i \leq r$ we find

$$\mathbb{P}(\mathcal{A}_i) \leq 2e^{-\frac{(2-i^2\varepsilon^2)^2}{8i\varepsilon} m^{3/2} n^{-1/2} + O(m^2 n^{-1})}$$

Note that Y_{τ_1} conditioned on $\mathcal{A}_i$ stochastically dominates a $\mathrm{Bin}(m((i-1)\varepsilon m^{1/2} n^{1/2} - m), p)$ distribution. Therefore,

$$\mathbb{P}(Y_{\tau_1} = 0 | \mathcal{A}_i) \leq (1-p)^{m((i-1)\varepsilon m^{1/2} n^{1/2} - m)} \leq e^{-(i-1)\varepsilon m^{3/2} n^{-1/2} + O(m^2 n^{-1})}$$

Combining the above and substituting into (2.20) yields

$$\mathbb{P}(Y_{\tau_1} = 0) \leq 2\sum_{i=1}^{r} e^{-((i-1)\varepsilon + \frac{(2-i^2\varepsilon^2)^2}{8i\varepsilon}) m^{3/2} n^{-1/2} + O(m^2 n^{-1})} + e^{-r\varepsilon m^{m/2} n^{-1/2} + O(m^2 n^{-1})}, \tag{2.21}$$
$$\leq (2r+1) e^{-(1+\gamma) m^{3/2} n^{-1/2} + O(m^2 n^{-1})}, \tag{2.22}$$

for some $\gamma > 0$ provided that ε is sufficiently small. The second term in (2.21) is a result of the fact $\mathbb{P}(A_{r+1}) \leq 1$. This proves the lemma and if one wishes for explicit constants, taking $\varepsilon = 0.025$, $r = 45$ works and gives $\beta < 100$, $\gamma > 0.06$. $\square$

The next stage in our proof is to show that a typical instance of $D(n, 1/n)$ has no

component of large excess and no exceptionally large components. This will allow us to use the bound from Lemma 2.2.3 to compute the expected number of large strongly connected components of $D(n, 1/n)$. The first result in this direction is an immediate corollary of a result of Łuczak and Seierstad [55].

Lemma 2.4.4 ([55]). *The probability that $D(n, 1/n)$ contains a strongly connected component of size at least $n^{1/3} \log\log n$ is $o_n(1)$.*

The next lemma ensures that there are not too many cycles which enables us to prove that the total excess is relatively small.

Lemma 2.4.5. *The probability that $D(n, p)$ contains more than $n^{1/6}$ cycles of length bounded above by $n^{1/3} \log\log(n)$ is $o_n(1)$.*

Proof. In this proof and subsequently we will use the convention that $\log^{(k)} x$ is the logarithm function composed with itself k times, while $(\log x)^k$ is its kth power. We shall show that the expected number of cycles of length at most $n^{1/3} \log^{(2)} n$ is $o(n^{1/6})$ at which point we may apply Markov's inequality. So let C be the random variable which counts the number of cycles of length at most $n^{1/3} \log^{(2)} n$ in $D(n, 1/n)$. We can calculate its expectation as

$$\mathbb{E}(C) = \sum_{k=1}^{n^{1/3} \log^{(2)} n} \binom{n}{k} \frac{k!}{k} p^k \leq \sum_{k=1}^{n^{1/3} \log^{(2)} n} \frac{1}{k} \tag{2.23}$$

We use the upper bound on the kth harmonic number $H_k \leq \log k + 1$, which allows us to deduce that

$$\mathbb{E}(C) \leq H_{n^{1/3} \log^{(2)} n} \leq \frac{1}{3} \log n + \log^{(3)} n + 1 \leq \log n = o(n^{1/6}). \tag{2.24}$$

Thus the lemma follows by Markov's inequality. □

Corollary 2.4.6. The probability that $D(n, 1/n)$ contains a component of excess at least $n^{1/6}$ and size at most $n^{1/3} \log\log n$ is $o_n(1)$.

Proof. If D is any strongly connected digraph with m vertices and excess k, then note that it must have at least $k+1$ cycles of length at most m. This can be seen by considering

the ear decomposition of D. The first ear must be a cycle, and each subsequent ear adds a path which must be contained in a cycle as D is strongly connected. So as we build the ear decomposition, each additional ear adds at least one cycle. As any ear decomposition of a strongly connected digraph of excess k has $k+1$ ears, then D must have at least $k+1$ cycles.

Thus, if D has k cycles, it must have excess at most $k-1$. So applying Lemma 2.4.5 completes the proof. □

Finally, we prove the main theorem of this section.

Proof of Theorem 2.4.1. Let $\mathcal{C}_1$ be the largest strongly connected component of $D(n, 1/n)$ and $L_1 = |\mathcal{C}_1|$. We want to compute $\mathbb{P}(L_1 \geq An^{1/3})$. Define the following three events,

$$\mathcal{E}_1 = \{L_1 \geq An^{1/3}\}$$
$$\mathcal{E}_2 = \{An^{1/3} \leq L_1 \leq n^{1/3} \log\log(n)\}$$
$$\mathcal{E}_3 = \{L_1 \geq n^{1/3} \log\log(n)\}$$

Clearly, $\mathcal{E}_1 \subseteq \mathcal{E}_2 \cup \mathcal{E}_3$ and by Lemma 2.4.4, $\mathbb{P}(\mathcal{E}_3) = o_n(1)$. If $\mathcal{F}$ is the event that $\mathcal{C}_1$ has excess at least $n^{1/6}$ then by Corollary 2.4.6, $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}) = o_n(1)$. All that remains is to give a bound on $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c)$. To this end let $N(A)$ be random variable which counts the number of strongly connected components of $D(n, 1/n)$ which have size between $An^{1/3}$ and $n^{1/3} \log\log n$ and excess bounded above by $n^{1/6}$. By Markov's inequality, we may deduce that $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c) \leq \mathbb{E}(N(A))$. Computing the expectation of $N(A)$,

$$\mathbb{E}(N(A)) = \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \sum_{k=0}^{n^{1/6}} \binom{n}{m} p^{m+k} Y(m,k) \mathbb{P}(Y_{\tau_1} = 0 | X_0 = m). \tag{2.25}$$

In Lemma 2.4.3 we showed that $\mathbb{P}(Y_{\tau_1} = 0 | X_0 = m) \leq \beta e^{-(1+\gamma)m^{3/2}n^{-1/2} + O(m^2 n^{-1})}$. Also,

using Lemma 2.2.3 we can check that

$$\sum_{k=0}^{n^{1/6}} Y(m,k)p^k \leq (m-1)! + C(m-1)!(m^3p)^{1/2}\sinh\big((m^3p)^{1/2}\big), \tag{2.26}$$

where the first term on the right hand side of (2.26) comes from the directed cycles and C is the same constant as in Lemma 2.2.3. As $\sinh(x) \leq e^x$ we can bound (2.26) by

$$\begin{aligned}\sum_{k=0}^{n^{1/6}} Y(m,k)p^k &\leq (m-1)!(1 + Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}}) \\ &\leq 2(m-1)!Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}}\end{aligned}$$

Combining these bounds and using $\binom{n}{m} \leq n^m/m!$ we deduce

$$\begin{aligned}\mathbb{E}(N(A)) &\leq \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \left(\frac{(np)^m}{m!}\right)\left(2(m-1)!Cm^{3/2}n^{-1/2}e^{m^{3/2}n^{-1/2}}\right)\left(\beta e^{-(1+\gamma)m^{3/2}n^{-1/2}+O(m^2n^{-1})}\right) \\ &= \sum_{m=An^{1/3}}^{n^{1/3}\log^2(n)} \frac{2\beta Cm^{1/2}}{n^{1/2}} e^{-\gamma m^{3/2}n^{-1/2}+O(m^2n^{-1})} \\ &\leq \int_{m=An^{1/3}}^{n^{1/3}\log^2(n)+1} \frac{2\beta Cm^{1/2}}{n^{1/2}} e^{-\frac{\gamma}{2}m^{3/2}n^{-1/2}}dm \end{aligned} \tag{2.27}$$

where (2.27) holds for all sufficiently large n. Now making the substitution $x = mn^{-1/3}$ we can remove the dependence of (2.27) on both m and n so that

$$\begin{aligned}\mathbb{E}(N(A)) &\leq 2\beta C \int_A^{\log^2(n)+n^{-1/3}} x^{1/2}e^{-\frac{\gamma}{2}x^{3/2}}dx \\ &\leq 2\beta C \int_A^{\infty} x^{1/2}e^{-\frac{\gamma}{2}x^{3/2}}dx \\ &= \frac{8\beta C}{3\gamma}\int_{\frac{\gamma A^{3/2}}{2}}^{\infty} e^{-t}dt = \frac{8\beta C}{3\gamma}e^{-\frac{\gamma A^{3/2}}{2}}\end{aligned} \tag{2.28}$$

So, by Markov's inequality $\mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c) \leq \zeta e^{-\eta A^{3/2}}$ where ζ and η are the corresponding

constants found in (2.28). So,

$$\mathbb{P}(L_1 \geq An^{1/3}) \leq \mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}^c) + \mathbb{P}(\mathcal{E}_2 \cap \mathcal{F}) + \mathbb{P}(\mathcal{E}_3) = \zeta e^{-\eta A^{3/2}} + o_n(1).$$

Calculating ζ and γ using the values for C, β and γ in Lemmas 2.2.3 and 2.4.3 yields $\zeta < 2 \times 10^7$ and $\eta > 0.03$. □

2.5 Adaptations for the Critical Window

In this section we sketch the adaptations one must make to the proofs of Theorems 2.3.1 and 2.4.1 such that they hold in the whole critical window, $p = n^{-1} + \lambda n^{-4/3}$ where $\lambda \in \mathbb{R}$.

2.5.1 Lower Bound

For Theorem 2.3.1, the adaptation is rather simple. We will still apply Janson's inequality and so we only need to recompute μ and Δ. Furthermore, the only difference in these calculations comes from replacing the term n^{-m-k} by p^{m+k}, and in fact the p^k in this turns out to make negligible changes. In this light, Lemma 2.3.3 changes to

Lemma 2.5.1.

$$\mathbb{E}(X) \geq \begin{cases} -e^{\frac{\lambda\delta}{2}} \log(\delta)/2 & \textit{if } \lambda \geq 0 \\ -e^{2\delta^{1/2}\lambda} \log(\delta)/2 & \textit{otherwise} \end{cases}$$

where the only difference in the proof is to bound $(1 + \lambda n^{-1/3})^m$ by its lowest value depending on whether $\lambda \geq 0$ or $\lambda < 0$. We bound this via

$$1 + x \geq \begin{cases} e^{\frac{x}{2}} & \text{if } 0 \leq x \leq 2 \\ e^{2x} & \text{if } -\frac{1}{2} \leq x \leq 0 \end{cases}$$

Furthermore, Lemma 2.3.5 changes to

Lemma 2.5.2. *For all sufficiently large n and small enough δ,*

$$\Delta \leq \begin{cases} e^{2\delta^{1/2}\lambda}\log(2) & \text{if } \lambda \geq 0 \\ e^{\delta\lambda}\log(2) & \text{otherwise} \end{cases}$$

The proof again is almost identical with the only change being to approximate the $(np)^m$ term. This time we seek an upper bound so use the approximation $1 + x \leq e^x$ which is valid for any x. We still need to split depending upon the sign of λ as for the above constants we upper bound $(np)^m$ by its largest possible value over the range $\delta n \leq m \leq 2\delta^{1/2}n$. Combining Lemmas 2.5.1 and 2.5.2 with the relevant constraints on δ in relation to λ yields Theorem 2.1.4.

2.5.2 Upper Bound

There is no significant (i.e. of order $e^{\lambda A}$) improvement which can be made with our current method of proof when $\lambda < 0$. This is because the gains we make computing the expectation in the proof of Theorem 2.4.1 are cancelled out by losses in the branching process considerations of Lemma 2.4.2.

When $\lambda > 0$ we cannot simply use our bound for $p = n^{-1}$ and thus an adaptation is necessary. Note that by monotonicity in p, the results of Lemmas 2.4.2 and 2.4.3 remain true for $p = n^{-1} + \lambda n^{-4/3}$ with $\lambda > 0$. The next adaptation which must be made is in equation (2.23) where now, the expectation becomes

$$\mathbb{E}(\mathcal{C}) \leq \sum_{k=1}^{n^{1/3}\log^{(2)} n} \frac{e^{k\lambda n^{-1/3}}}{k} \leq \sum_{k=1}^{n^{1/3}\log^{(2)} n} \frac{(\log n)^{\lambda}}{k} \leq 2(\log n)^{\lambda+1} = o(n^{1/6})$$

Thus allowing us to deduce the result of Corollary 2.4.6 as before. Finally all that remains is to conclude the proof of Theorem 2.1.5. Ignoring lower order terms, the only difference to the proof compared to that of Theorem 2.4.1 is in the computation of $\mathbb{E}(N(A))$ where

we must change the term $(np)^m$. Thus the integral in (2.27) becomes

$$\int_{m=An^{1/3}}^{n^{1/3}\log^{(2)} n+1} \frac{2\beta C m^{1/2}}{n^{1/2}} e^{-\frac{\gamma}{2}m^{3/2}n^{-1/2}+\lambda m n^{-1/3}} dm \tag{2.29}$$

This is much more complex than before due to the extra term in the exponent. However we are still able to give a bound after making the obvious substitution $t = \frac{\gamma}{2}m^{3/2}n^{-1/2} - \lambda m n^{-1/3}$, we obtain

$$\begin{aligned}\mathbb{E}(N(A)) &\leq \frac{8\beta C}{3\gamma}\int_{\frac{\gamma}{2}A^{3/2}-\lambda A}^{\infty} \frac{m^{1/2}n^{-1/2}}{m^{1/2}n^{-1/2} - \frac{4\lambda n^{-1/3}}{3\gamma}} e^{-t} dt \\ &\leq \frac{10\beta C}{3\gamma}\int_{\frac{\gamma}{2}A^{3/2}-\lambda A}^{\infty} e^{-t} dt = \frac{10\beta C}{3\gamma} e^{-\frac{\gamma}{2}A^{3/2}+\lambda A}\end{aligned} \tag{2.30}$$

which is of the claimed form. Note the second inequality holds for A sufficiently large compared to λ.

2.6 Concluding Remarks

In this chapter we have proven that inside the critical window, $p = n^{-1} + \lambda n^{-4/3}$, the largest component of $D(n,p)$ has size $\Theta_p(n^{1/3})$. Furthermore, we have given bounds on the tail probabilities of the distribution of the size of the largest component. Combining this result with previous work of Karp [47] and Łuczak [53] allows us to deduce that $D(n,p)$ exhibits a "double-jump" phenomenon at the point $p = n^{-1}$. However, there are still a large number of open questions regarding the giant component in $D(n,p)$. Recently, Goldschmidt and Stephenson [31] found a scaling limit for the sizes of all the strong components in the critical random digraph. This scaling limit is a little difficult to work with directly and so it would be interesting to know if there is a more explicit form for the size of the largest component. In $G(n,p)$ such a result was given by Pittel [69]

Question 1. Is there an explicit description of the limiting distribution of the largest component of $D(n,p)$?

Given the strong connection between $G(n,p)$ and $D(n,p)$, it seems likely that the limit distributions, $X^\lambda = n^{-2/3}|\mathcal{C}_1(G(n,p))|$ and $Y^\lambda = n^{-1/3}|\mathcal{C}_1(D(n,p))|$ (where $p = n^{-1} + \lambda n^{-4/3}$) are closely related. For larger p, previous work [47, 54] has found that the size of the giant strongly connected component in $D(n,p)$ is related to the size of the square of the giant component in $G(n,p)$. That is, if $|\mathcal{C}_1(G(n,p)| \sim \alpha(n)n$, then $|\mathcal{C}_1(D(n,p)| \sim \alpha(n)^2 n$. Note that the result found in Theorem 2.1.5 is consistent with this pattern as here we have an exponent of order $A^{3/2}$ while for $G(n,p)$ a similar result is true with exponent A^3 implying that the probability we find a component of size $Bn^{2/3}$ in $G(n,p)$ is similar to the probability of finding a component of size $B^2 n^{1/3}$ in $D(n,p)$ (assuming both bounds are close to tight). As such, we make the following conjecture to explain this pattern.

Conjecture 2.6.1. If X^λ and Y^λ are the distributions defined above and X_1^λ, X_2^λ are independent copies of X^λ then, $Y^\lambda = X_1^\lambda X_2^\lambda$.

Finally, we consider the transitive closure of random digraphs. The *transitive closure* of a digraph D is $cl(D)$ a digraph on the same vertex set as D and such that uv is an edge of $cl(D)$ if and only if there is a directed path from u to v in D. Equivalently, $cl(D)$ is the smallest digraph containing D such that the relation R defined by uRv if and only if uv is an edge is transitive. Karp [47] gave a linear time algorithm to compute the transitive closure of a digraph from the model $D(n,p)$ provided that $p \leq (1-\varepsilon)n^{-1}$ or $p \geq (1+\varepsilon)n^{-1}$. For all other p this algorithm runs in time $O(f(n)(n \log n)^{4/3})$ where $f(n)$ is any $\omega(1)$ function. Now that we know more about the structure of $D(n,p)$ for p close to n^{-1}, it may be possible to adapt Karp's algorithm and obtain a better time complexity.

Question 2. Does there exist a linear time algorithm to compute the transitive closure of $D(n,p)$ when $(1-\varepsilon)n^{-1} \leq p \leq (1+\varepsilon)n^{-1}$?

CHAPTER 3

BARELY SUBCRITICAL GRAPHS FROM THE CONFIGURATION MODEL

3.1 Introduction

Let $[n] := \{1, \ldots, n\}$ be a set of n vertices. Let $\mathbf{d}_n = (d_1, \ldots, d_n)$ be a degree sequence with $m := \sum_{i \in [n]} d_i$ an even positive integer. Without loss of generality, we will assume that $d_1 \leq \cdots \leq d_n$. Additionally, we may assume that $d_1 \geq 1$; if there are elements with degree 0 we can remove them and study the remainder sequence. Let $\Delta = \Delta_n$ be the maximum degree of $\mathbf{d}_n$.

The *configuration model*, denoted by $\mathbb{CM}_n = \mathbb{CM}_n(\mathbf{d}_n)$, is the random multigraph on $[n]$ generated by giving d_i *half-edges* (or stubs) to vertex i, and then pairing the half-edges uniformly at random. The *uniform model*, denoted by $\mathbb{G}_n = \mathbb{G}_n(\mathbf{d}_n)$, is the random simple graph on $[n]$ obtained by choosing a simple graph uniformly at random among all graphs on $[n]$ where vertex i has degree d_i. Throughout this chapter, all the results on the uniform model will assume that the sequence $\mathbf{d}_n$ is *graphical*; that is, there exists at least one graph on $[n]$ with such degree sequence.

For any graph G on $[n]$, let $L_1(G)$ denote the order of a largest component. A central problem in random graph theory is to find a parameter of the model α such that L_1 undergoes a phase transition at $\alpha = \alpha_0$. The set of parameters is then divided into subcritical ($\alpha < \alpha_0$), critical ($\alpha = (1 + o(1))\alpha_0$) and supercritical ($\alpha > \alpha_0$). A further

problem is the study the critical window: that is, to find parameters β^-, β^+, such that L_1 behaves essentially the same for any $\alpha \in (\alpha_0 - \beta^-, \alpha_0 + \beta^+)$ and the critical region is further divided into barely subcritical, $\alpha_0 - \alpha > \beta^-$ and barely supercritical, where $\alpha - \alpha_0 > \beta^+$.

The main goal of this chapter is to study the largest component phase transition $L_1(\mathbb{CM}_n)$ and $L_1(\mathbb{G}_n)$ in the subcritical and the barely subcritical regimes. Generally speaking, the study of configuration model is simpler due to the existence of a explicit model with good independence properties. In contrast, most of the results existing for $\mathbb{G}_n$ arise from $\mathbb{CM}_n$ by observing that the probability that $\mathbb{CM}_n$ generates a simple graph is sufficiently large.

In order to understand the phase transition, define

$$Q = Q_n(\mathbf{d}_n) := \frac{1}{m} \sum_{i \in [n]} d_i(d_i - 2) \ , \tag{3.1}$$

$$R = R_n(\mathbf{d}_n) := \frac{1}{m} \sum_{i \in [n]} d_i(d_i - 2)^2 \ . \tag{3.2}$$

In the first part of the chapter, we will focus on the case $Q_n \leq 0$. It is easy to check that the bound on Q_n implies $\Delta_n = O(\sqrt{n})$ and $m \leq 2n$. Also note the implicit bound on the maximum degree $\Delta_n = O(n^{1/3} R_n^{1/3})$ obtained by just considering the contribution of a vertex of maximum degree to R_n.

Let D_n be the degree of a uniform random vertex and let $\hat{D}_n$ be its size-biased distribution; that is, for $k \geq 1$,

$$\mathbb{P}\left(\hat{D}_n = k\right) = \frac{k\mathbb{P}\left(D_n = k\right)}{m} \ . \tag{3.3}$$

For $b, h \in \mathbb{N}$, let $\mathcal{L}(b, h) := \{b + hk : k \in \mathbb{Z}\}$ be the integer lattice containing b with step h. Let h_n be the largest integer h such that $\mathbb{P}(D_n \in \mathcal{L}(b, h)) = 1$ for some $b \in \mathbb{N}$.

We will study the $\mathbb{CM}_n$ under the following mild conditions on the degree sequence:

Assumption 3.1.1. *There exists a discrete random variable D supported on $\mathbb{Z}_{\geq 0}$ such*

that

(i) $D_n \to D$ *in distribution;*

(ii) $Q_n \to 0$*;*

(iii) $\mathbb{P}(D \notin \{0,2\}) > 0$*;*

(iv) if h *is the largest integer such that* $\mathbb{P}(D \in \mathcal{L}(b,h))$ *for some* $b \in \mathbb{N}$*, then* $h_n = h$ *for all* n*.*

(v) $\mathbb{E}(D_n^4) \leq \Delta_n^{1/2}$

Remark 3.1.2. Conditions (i)-(iii) are usual in this setting. In particular, they imply that R_n is bounded away from zero, which will be often used in the proofs.

Condition (iv) simply asks that the limiting degree distribution D has the same step as the random variables D_n which converge to it. This restriction is not particularly strong and forbids no limiting degree sequence, only the way in which we converge to it.

Condition (v) is the most restrictive one. As $Q_n = o(1)$, we have $\mathbb{E}[D_n^2] = O(1)$, which implies $\mathbb{E}[D_n^4] = O(\Delta_n^2)$. Thus, this condition can be understood as a "polynomial limitation" on the contribution of large degree vertices to the fourth moment. It would be interesting to see up to which point a condition on the fourth moment is needed.

Our first result upper bounds the size of the largest component when Q is not too large with respect to R.

Theorem 3.1.3. Let $\epsilon > 0$. Let $\mathbf{d}_n$ be a degree sequence satisfying Assumption 3.1.1 and $\Delta|Q| = o(R)$. If $Q \leq -\omega(n)n^{-1/3}R^{2/3}$ for some $\omega(n) \to \infty$, then

$$\mathbb{P}\left(L_1(\mathbb{CM}_n(\mathbf{d}_n)) \leq (1+\epsilon)\frac{2R}{Q^2}\log\left(\frac{|Q|^3 n}{R^2}\right)\right) = 1 - o(1)\ . \tag{3.4}$$

Remark 3.1.4. As noted by [35], under the condition $|Q|\Delta = o(R)$ the critical window is $|Q| = O(n^{-1/3}R^{2/3})$. Therefore, Theorem 3.1.3 bounds the largest component in the whole barely subcritical regime. See Section 3.1.1 for further discussion.

Let $\eta := \eta_n = \hat{D}_n - 2$ and consider its moment generating function

$$\varphi(\theta) := \varphi_n(\theta) = \mathbb{E}\left[e^{\theta\eta_n}\right] \ . \tag{3.5}$$

Theorem 3.1.3 is in fact a consequence of a more general result that does not require a bound of Q in terms of R.

Theorem 3.1.5. Let $\epsilon > 0$. Let $\mathbf{d}_n$ be a degree sequence satisfying Assumption 3.1.1 and $\Delta_n \leq n^{1/6}$. Let $\theta_0 \in (0,1)$ be the smallest solution θ of $\varphi'(\theta) = 0$. Define

$$T_n := \frac{1}{\log(\varphi(\theta_0))^{-1}} \cdot \log\left(\frac{(\log(\varphi(\theta_0))^{-1})^{3/2}}{\varphi''(\theta_0)^{1/2}}\mathbb{E}\left[D_n e^{\theta_0 D_n}\right] n\right) \ . \tag{3.6}$$

If $\theta_0 m \geq \omega(n) T_n$ for some $\omega(n) \to \infty$, then

$$\mathbb{P}\left(L_1(\mathbb{CM}_n(\mathbf{d}_n)) \leq (1+\epsilon)T_n\right) = 1 - o(1) \ . \tag{3.7}$$

Remark 3.1.6. The value θ_0 exists and is bounded as $n \to \infty$. We have $\varphi(0) = 1$ and $\varphi'(0) = Q < 0$. Recall that η is supported in $\{-1, 0, 1, \dots\}$. By Assumption 3.1.1, $\{D \geq 3\}$ happens with positive probability and so if we define $p := \mathbb{P}(D \geq 3)$, then $\mathbb{P}(\eta \geq 1) \geq p$ and $\varphi(\theta) \geq (1-p)e^{-\theta} + pe^{\theta}$. So $\varphi(\theta) \to \infty$ as $\theta \to \infty$ and it must at some point have positive derivative. Thus, there exists θ_0 such that $\varphi'(\theta_0) = 0$.

It is interesting to understand if these results also hold in the uniform setting. One can use the following result to transfer from $\mathbb{CM}_n$ to $\mathbb{G}_n$.

Theorem 3.1.7 (Janson [41])**.** Let $\mathbf{d}_n$ be a degree sequence satisfying $m = \Theta(n)$ and $\mathbb{E}\left[D_n^2\right] = O(1)$. Then

$$\mathbb{P}\left(\mathbb{CM}_n(\mathbf{d}_n) \text{ simple}\right) = \exp\left(-\frac{1}{m}\sum_{i\in[n]} d_i^2\right) > 0 \ , \tag{3.8}$$

and conditioned on being simple, $\mathbb{CM}_n$ has the same law as $\mathbb{G}_n$. Therefore, any result

that holds with probability $1-o(1)$ for $\mathbb{CM}_n(\mathbf{d}_n)$, also holds with probability $1-o(1)$ for $\mathbb{G}_n(\mathbf{d}_n)$.

As Theorem 3.1.3 and Theorem 3.1.5 assume that $Q_n \leq 0$, we have that $\mathbb{E}\left[D_n^2\right] = O(1)$ and we can use Theorem 3.1.7 to transfer their conclusions to $\mathbb{G}_n$, provided that their hypothesis are satisfied.

The second part of our chapter focuses on the size of the largest component in the barely subcritical regime of $\mathbb{G}_n$ without further assumptions on the degree sequence. The lack of a tractable model for $\mathbb{G}_n$ hampers its analysis and the upper bounds obtained are weaker than the ones obtained for $\mathbb{CM}_n$ and probably not of the right order.

Let S_* be a smallest set of vertices of largest degree that satisfies

$$\sum_{u\in[n]\setminus S_*} d_u(d_u-2) \leq 0 \tag{3.9}$$

and define

$$m_* = \sum_{v\in S_*} d_v \ . \tag{3.10}$$

In particular, if $Q \leq 0$, then $S_* = \emptyset$ and $m_* = 0$.

For any $m_0 \geq 0$ and $Q_0 \leq 0$, we call $\mathbf{d}_n$ an (m_0, Q_0)-*subcritical* degree sequence if there exists $S \subseteq [n]$ with $\sum_{v\in S} d_v \leq m_0$ and

$$\frac{1}{m}\sum_{w\in[n]\setminus S} d_w(d_w-2) \leq Q_0 \ .$$

Our most general result on $\mathbb{G}_n$ is the following.

Theorem 3.1.8. Let $\mathbf{d}_n$ be an (m_0, Q_0)-subcritical degree sequences for some parameters satisfying $m_0 \geq 3m_*$, $m_0|Q_0| \geq (\Delta|Q_0|+R)\log\left(\frac{nQ_0^2}{\Delta|Q_0|+R}\right)$ and $Q_0^2 n \geq \omega(n)m_0$ for some

$\omega(n) \to \infty$. Then,

$$\partial L_1(\mathbb{G}_n(\mathbf{d}_n)) = O(m_0/|Q_0|) = 1 - o(1) \ .$$

Remark 3.1.9 (Infinite degree variance)**.** The main strength of Theorem 3.1.8 is that it applies to degree sequences with subcritical behaviour but infinite degree variance. To our knowledge, the only results available in this setting are of the form $L_1(\mathbb{G}_n(\mathbf{d}_n)) = o(n)$ [8, 45]. Note that, even if such whp results were available for $\mathbb{CM}_n$, Theorem 3.1.7 is not strong enough to transfer them to $\mathbb{G}_n$.

Remark 3.1.10 (The $Q \leq 0$ case)**.** To compare it with previous work, let us get more explicit results for the case $Q \leq 0$ (i.e. $\mathbb{E}[D_n^2] \leq 2\mathbb{E}[D_n]$). In this case, $m_* = 0$ and we can choose $Q_0 := Q$ and $m_0 := (\Delta + R/|Q|) \log\left(\frac{nQ_0^2}{\Delta|Q_0|+R}\right)$. Also note that $Q \leq 0$ implies $R = O(\Delta)$.

(a) If $|Q|$ is bounded away from zero, then all conditions in Theorem 3.1.8 are satisfied and

$$L_1(\mathbb{G}_n(\mathbf{d}_n)) = O(\Delta \log n) \ . \tag{3.11}$$

(b) If $|Q| = o(1)$, then we split depending on how R and $\Delta|Q|$ compare to each other.

(b.1) If $\Delta|Q| = O(R)$, then for any $Q \leq -\omega(n)n^{-1/3}R^{1/3}$,

$$L_1(\mathbb{G}_n(\mathbf{d}_n)) = O\left(\frac{R}{Q^2}\log\left(\frac{nQ^2}{R}\right)\right) \ , \tag{3.12}$$

obtaining a weaker version of Theorem 3.1.3 under no additional assumptions.

(b.2) If $R = O(\Delta|Q|)$, then for any $Q \leq -\omega(n)n^{-1/2}\Delta^{1/2}$,

$$L_1(\mathbb{G}_n(\mathbf{d}_n)) = O\left(\frac{\Delta}{|Q|}\log\left(\frac{nQ}{\Delta}\right)\right) \ . \tag{3.13}$$

Remark 3.1.11. If $Q_0^2 n = O(m_0)$, then the behaviour of $\mathbb{G}_n$ is no longer (barely) subcritical. It is interesting to study the size of the largest component in this case.

We finally provide the existence of infinitely many degree sequences that show the tightness of some of our upper bounds.

Proposition 3.1.12. For any $Q < 0$, $\Delta = o(\sqrt{n})$ and $\log n = o(\Delta)$, there exists a degree sequence $\widehat{\mathbf{d}}_n$ with $\Delta_n(\widehat{\mathbf{d}}_n) = \Delta$, $Q_n(\widehat{\mathbf{d}}_n) \sim Q$ and $R = R_n(\widehat{\mathbf{d}}_n) \sim \Delta$, such that

$$\partial L_1(\mathbb{G}_n(\widehat{\mathbf{d}}_n)) \geq (1 + o(1)) \frac{2R}{Q^2} \log\left(\frac{n}{R^2}\right) = 1 - o(1) \ .$$

Remark 3.1.13. We can compare the lower bound in Proposition 3.1.12 with our upper bounds. The degree sequence $\widehat{\mathbf{d}}_n$ satisfies $R \sim \Delta$, so $\Delta|Q| = O(R)$. In the case $\Delta < n^{1/2-\delta}$ for some constant $\delta > 0$, the proposition gives a family of degree sequences for which Equation (3.11) is of the right order.

While Proposition 3.1.12 is only stated for Q bounded away from zero, one could similarly define degree sequences $\widehat{\mathbf{d}}_n$ for which $Q = o(1)$, in which case $\Delta|Q| = o(R)$. Provided that $Q \leq -\omega(n) n^{-1/3} R^{2/3}$ for some $\omega(n) \to \infty$, one can obtain the lower bound in Equation (3.90) that coincides asymptotically with Theorem 3.1.3 and, up to logarithmic terms, with Equation (3.12). (See Remark 3.5.3.)

3.1.1 Previous work

The foundational paper of Erdős and Rényi [24] located the phase transition for the existence of a linear order component in a uniformly chosen graph on n vertices and m edges, $\mathbb{G}(n, m)$, showing that the order of the largest component undergoes a double jump at $m = n/2$, in particular $L_1(\mathbb{G}(n, m)) = O(\log n)$ if $m \leq cn$ and $c < 1/2$, $L_1(\mathbb{G}(n, m)) = \Theta(n^{2/3})$ if $m = n/2$, and $L_1(\mathbb{G}(n, m)) = \Theta(n)$ if $m \leq cn$ and $c > 1/2$. This result can be easily transferred to the Binomial random graph $\mathbb{G}(n, p)$ with $p = 2m/n$, which has become the reference model for random graphs. The size of the largest component in all regimes is well understood, see e.g. Sections 4 and 5 in [36].

The study of the phase transitions in random graphs with given degree sequences was pioneered by Molloy and Reed [58]. The so-called *Molloy-Reed criterion* determines the phase transition at $Q = 0$, provided that the degree sequence satisfies a number of technical conditions. The criterion has been extended to degree sequences with bounded degree variance [43] and uniformly integrable sequences [8], providing the asymptotic value of L_1 in the supercritical regime $Q > 0$ in terms of the survival probability of a branching process, similarly as it is $\mathbb{G}(n, p)$ case. Interestingly, the criterion is no longer valid for general sequences of graphs due to the presence of high degree vertices (hubs) or an extremely large number of degree 2 vertices. Joos et al. [45] gave an extended criterion that determines whether any given sequence typically provides a linear order component.

While the behaviour of the largest component in the supercritical regime resembles the simpler Erdős-Rényi model, this does not happen in the subcritical one, when $Q < 0$. Trivially, we have $L_1(\mathbb{G}_n) \geq \Delta + 1$ which could be much larger than logarithmic. In [58] the authors showed that $L_1(\mathbb{CM}_n) = O(\Delta^2 \log n)$ for subcritical sequences. More precise results are known for power-law degree sequences. Durrett [21] conjectured[1] that if $\mathbb{P}(D_n = k) \sim ck^{-\gamma}$ for some $\gamma > 3$ and $c > 0$, then $L_1(\mathbb{CM}_n) = O(\Delta)$. In this setting, $\gamma > 3$ implies $\mathbb{E}[D_n^{\gamma-1}] = O(1)$. Pittel [70] showed that $L_1(\mathbb{CM}_n) = O(\Delta \log n)$ for subpower-law distributions. Janson [40] proved a strong version of the conjecture: if $\mathbb{P}(D_n \geq k) = O(k^{1-\gamma})$ for some $\gamma > 3$, then

$$L_1(\mathbb{CM}_n) = \frac{\Delta}{|Q|} + o(n^{1/(\gamma-1)}) \tag{3.14}$$

For power-law distributions, whp we have $\Delta = \Theta(n^{1/(\gamma-1)})$, and the second term is negligible. From the intuitive point of view, the largest component is obtained by starting a subcritical branching process with expected offspring $1 + Q$ from each vertex adjacent to the vertex of largest degree. The expected total progeny of such process is $1/|Q|$. One can interpret the result of Theorem 3.1.8 in a similar spirit: in the largest component

[1] In fact, this was conjectured for a slightly different model where the degrees are i.i.d. copies of D_n conditioned on their sum being even.

there might be at most $O(m_0)$ edges and from each of these edges a piece of size $O(1/|Q_0|)$ hangs. For $Q < 0$, (3.14) can be compared to the weaker bound Equation (3.11) that holds regardless of the shape of the degree sequence tail. As shown in Remark 3.1.10, for general degree sequences Equation (3.11) cannot be improved.

The critical regime has attracted a lot of interest in recent years [17, 18, 35, 37, 46, 72] with several papers specialising on the finite second or finite third moment cases. Here we focus on the results known for the barely subcritical regime.

Riordan [72] showed that if $\Delta = O(1)$ then Equation (3.4) holds, even more, one has asymptotic equality and control on the second order term. Theorem 3.1.5 can be seen as a generalization of the upper bound in [72] to a wider class of sequences that allows $\Delta \to \infty$ as $n \to \infty$.

Hatami and Molloy [35] studied the critical window under some mild conditions on the degree sequence. They showed that $|Q| = O(n^{-1/3}R^{2/3})$ is the critical window of $\mathbb{CM}_n$. Regarding the barely subcritical regime, for $Q \leq -\omega(n)n^{-1/3}R^{2/3}$ with $\omega(n) \to \infty$, they showed that whp

$$L_1(\mathbb{CM}_n) = O\left(\sqrt{\frac{n}{|Q|}}\right) . \tag{3.15}$$

One can check that Equation (3.15) coincides in order with Equation (3.4) at the boundary of the critical window $|Q| = \Theta(n^{-1/3}R^{2/3})$, while Equation (3.4) improves Equation (3.15) in the whole barely subcritical regime, provided that Assumption 3.1.1 holds.

Under infinite variance, the probability of $\mathbb{CM}_n$ being simple can be exponentially small in n. Thus, only results that hold with exponentially high probability can be transferred from $\mathbb{CM}_n$ to $\mathbb{G}_n$, see e.g. [8]. Another approach is to study $\mathbb{G}_n$ directly using the *switching method* [45]. In both cases, the best bound given in the subcritical regime is $L_1(\mathbb{G}_n) = o(n)$. Theorem 3.1.8 provides the first explicit general bound to L_1 at subcriticallity for infinite variance degree sequences. As discussed in Remark 3.1.13, this bound cannot be substantially improved without further assumptions.

It thus remains as an open question to determine the exact size of the largest component in the (barely) subcritical regime. Hofstad, Janson and Łuczak conjectured that $L_1(\mathbb{CM}_n)$ is concentrated in this regime [37]. Supported by the result of Riordan for constant maximum degree, we conjecture that the upper bound in Equation (3.7) is asymptotically tight for all degree sequences that satisfy some mild assumption. Note that certain condition on the degree sequence is needed, as for some particular subcritical degree sequences $L_1(\mathbb{CM}_n)$ is non-concentrated (see Remark 3.5.2).

3.2 A local limit theorem

A local limit theorem estimates the probability distribution of a suitably rescaled sum of independent random variables, by the density function of a Gaussian random variable. Local limit theorems are a useful tool to determine the component size in random graphs [63, 72]. For our application, we will need the step distribution to allow for the existence of very large degree, as well as the fact that the degree sequence may be supported on an lattice with step different than 1. This prevents us from using classical results such as Berry-Esseen Theorem (see [22, Theorem 3.4.9]). Our goal is to develop a very precise local limit theorem which will allow us to deal with our step distributions. Our result is based on previous local limit theorems by Doney [20] and Mukhin [61, 60] from which we derive more explicit error bounds. In particular the main result of this section is following,

Theorem 3.2.1. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables taking values on $\mathcal{L}(v_0, h)$. Define $S_n = \sum_{i=1}^n X_i$. Suppose that $\mu = \mathbb{E}(X_1) = 0$, $\sigma^2 = \mathrm{Var}(X_1)$ and $\gamma = \mathbb{E}|X_1|^3$, and let $\varphi(t)$ be the characteristic function of X_1. Then,

$$\sup_{w \in \mathcal{L}(nv_0,h)} \left| \mathbb{P}(S_n = w) - \frac{h}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{w^2}{2n\sigma^2}\right) \right| \leq \frac{32h\gamma}{\sigma^4 n} + \frac{h}{\pi} \int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt. \tag{3.16}$$

To prove this theorem, we will require the following Fourier inverse theorems which

can be found in the book of Durett [22].

Theorem 3.2.2 (Continuous Fourier Inverse Theorem)**.** Suppose that X is a random variable with characteristic function $\varphi_X(t)$. Suppose further that $\varphi_X(t)$ is integrable, then X is a continuous random variable with density function $f(y)$ defined by

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ity} \varphi_X(t) dt.$$

Theorem 3.2.3 (Discrete Fourier Inverse Theorem)**.** Let X be a random variable with characteristic function $\varphi_X(t)$. Suppose that there exists $h > 0$ such that $\mathbb{P}(X \in h\mathbb{Z}) = 1$. Then for any $a \in h\mathbb{Z}$,

$$\mathbb{P}(X = a) = \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-ita} \varphi_X(t) dt.$$

Finally, as we are interested in local limit theorems it will be useful to note that the characteristic function of the standard normal distribution is given by $N(t) = e^{-\frac{t^2}{2}}$.

Proof of Theorem 3.2.1. Let $\varphi(t)$ be the characteristic function of X_1 and $\psi_n(t)$ the characteristic function of S_n. By basic properties of characteristic functions, it is easy to see that $\psi_n(t) = \varphi(t)^n$. By Theorems 3.2.2 and 3.2.3 we may deduce that

$$\mathbb{P}(S_n = w) - \frac{h}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{w^2}{2n\sigma^2}\right) = \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} e^{-itw} (\psi_n(t) - N(t\sigma\sqrt{n})) dt - \frac{h}{\pi} \int_{\frac{\pi}{h}}^{\infty} e^{-itw} N(t\sigma\sqrt{n}) dt.$$

Therefore, by applying various forms of the triangle law we obtain the bound

$$\left| \mathbb{P}(S_n = w) - \frac{h}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{w^2}{2n\sigma^2}\right) \right| \leq \frac{h}{2\pi} \int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |\psi_n(t) - N(t\sigma\sqrt{n})| dt + \frac{h}{\pi} \int_{\frac{\pi}{h}}^{\infty} N(t\sigma\sqrt{n}) dt. \tag{3.17}$$

To bound the first integral in (3.17) we split it into three parts. For $\varepsilon > 0$ (which we shall pick later) we have

$$\int_{-\frac{\pi}{h}}^{\frac{\pi}{h}} |\psi_n(t) - N(t\sigma\sqrt{n})| dt \leq \int_{-\varepsilon}^{\varepsilon} |\psi_n(t) - N(t\sigma\sqrt{n})| dt + 2 \int_{\varepsilon}^{\frac{\pi}{h}} N(t\sigma\sqrt{n}) dt + 2 \int_{\varepsilon}^{\frac{\pi}{h}} |\varphi(t)|^n dt \tag{3.18}$$

The idea of this bound being that both $\psi_n(t)$ and $N(t\sigma\sqrt{n})$ only contribute a non-trivial amount to the left hand side of (3.18) for t very close to 0. In bounding the first integral of (3.18) we will use the following lemma [67, Page 109, Lemma 1].

Lemma 3.2.4. *Let $X_1, \ldots, X_n$ be independent random variables with $\mathbb{E}(X_i) = 0$, $\sigma_i^2 = \mathrm{Var}(X_i)$ and $\gamma_i = \mathbb{E}|X_i|^3$. Define*

$$B_n := \sum_{i=1}^{n} \sigma_i^2 \qquad L_n := \frac{\sum_{i=1}^{n} \gamma_i}{B_n^{3/2}} \qquad T_n := \frac{\sum_{i=1}^{n} X_i}{B_n^{1/2}}$$

Let $f_n(t)$ be the characteristic function of T_n. Then,

$$|f_n(t) - e^{-\frac{t^2}{2}}| \leq 16 L_n |t|^3 e^{-\frac{t^2}{3}} \quad \textit{for} \quad |t| \leq \frac{1}{4L_n}. \tag{3.19}$$

Clearly this is applicable in our setting, however we need to rescale first which allows us to deduce

$$|\psi_n(t) - N(t\sigma\sqrt{n})| \leq 16\gamma n |t|^3 e^{-\frac{t^2\sigma^2 n}{3}} \quad \text{for} \quad |t| \leq \frac{\sigma^2}{4\gamma}.$$

So, for any $\varepsilon \leq \sigma^2/(4\gamma)$ we have

$$\begin{aligned} \int_{-\varepsilon}^{\varepsilon} |\psi_n(t) - N(t\sigma\sqrt{n})| dt &\leq 16\gamma n \int_{-\varepsilon}^{\varepsilon} |t|^3 e^{-\frac{t^2\sigma^2 n}{3}} dt \\ &\leq 16\gamma n \int_{-\infty}^{\infty} |t|^3 e^{-\frac{t^2\sigma^2 n}{3}} dt \\ &= \frac{16\gamma}{\sigma^4 n} \int_{-\infty}^{\infty} |t|^3 e^{-\frac{t^2}{3}} dt = \frac{144\gamma}{\sigma^4 n}. \end{aligned} \tag{3.20}$$

The next step is to bound the second term of (3.18). Note that we can combine this with bounding the second term of (3.17), so we require to give an upper bound on

$$\int_{\varepsilon}^{\infty} N(t\sigma\sqrt{n}) dt = \frac{1}{\sigma\sqrt{n}} \int_{\varepsilon\sigma\sqrt{n}}^{\infty} e^{-\frac{t^2}{2}} dt = \frac{\sqrt{2\pi}}{\sigma\sqrt{n}} \mathbb{P}(\mathcal{N}(0,1) > \varepsilon\sigma\sqrt{n}). \tag{3.21}$$

We use the Chernoff's bound for the standard normal distribution, $\mathbb{P}(\mathcal{N}(0,1) > x) \leq e^{-\frac{x^2}{2}}$

and the simple inequality $e^{-x} \leq x^{-1/2}$ for $x > 0$, to obtain

$$\int_{\varepsilon}^{\infty} N(t\sigma\sqrt{n})dt = \frac{\sqrt{2\pi}}{\sigma\sqrt{n}}\mathbb{P}(\mathcal{N}(0,1) > \varepsilon\sigma\sqrt{n}) \leq \frac{2\sqrt{\pi}}{\varepsilon\sigma^2 n} \tag{3.22}$$

Choosing $\varepsilon = \sigma^2/(4\gamma)$ and combining (3.17), (3.18), (3.20) and (3.22), we find that for any $w \in \mathcal{L}(nv_0, h)$,

$$\begin{aligned}\left|\mathbb{P}(S_n = w) - \frac{h}{\sqrt{2\pi n\sigma^2}}\exp\left(-\frac{w^2}{2n\sigma^2}\right)\right| &\leq \left(\frac{72}{\pi} + \frac{16}{\sqrt{\pi}}\right)\frac{h\gamma}{\sigma^4 n} + \frac{h}{\pi}\int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt \\ &\leq \frac{32h\gamma}{\sigma^4 n} + \frac{h}{\pi}\int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt. \end{aligned} \tag{3.23}$$

Concluding the proof of the theorem. □

For the remainder of this section we will focus on bounding the integral term in the RHS of (3.16). To this end we introduce the parameter $H_D(X)$, which generalises a similar parameter introduced by Mukhin [60, 61]. For a real-valued random variable X, we define $X^* = X - X'$ to be the *symmetrisation* of X, where X' is an independent copy of X. Furthermore, for $\alpha \in \mathbb{R}$ define $\langle\alpha\rangle$ to be the distance from α to the nearest integer. Then for a random variable X and $d \in \mathbb{R}$ we define the following parameters

$$H(X, d) := \mathbb{E}\langle X^* d\rangle^2.$$

The parameter $H(X, d)$ measures in a certain sense how close is X^* to be a random variable supported on a lattice with step $1/|d|$.

The following lemma from [60] will be useful.

Lemma 3.2.5. *If $\varphi(t)$ is the characteristic function of the random variable X then*

$$4H\left(X, \frac{t}{2\pi}\right) \leq 1 - |\varphi(t)| \leq 2\pi^2 H\left(X, \frac{t}{2\pi}\right) \tag{3.24}$$

We provide a full proof of the statement, for the sake of completeness.

Proof. We look at the characteristic function of X^*, $\varphi^*(t)$. Note that X^* is by definition symmetric around the origin and hence so is $\varphi^*(t)$. Writing $\mathcal{D}(X^*)$ for the domain of X^* which is discrete, we have

$$\varphi^*(t) = \frac{\varphi^*(t) + \varphi^*(-t)}{2} = \sum_{x \in \mathcal{D}(X^*)} \frac{e^{itx} + e^{-itx}}{2} \mathbb{P}(X^* = x) = \sum_{x \in \mathcal{D}(X^*)} \cos(tx) \mathbb{P}(X^* = x). \tag{3.25}$$

As $\cos(x)$ is symmetric around π and periodic with period 2π, we have the identity

$$\cos(x) = \cos\left(2\pi \left\langle \frac{x}{2\pi} \right\rangle\right).$$

We can use this identity to rewrite (3.25) as

$$\varphi^*(t) = \sum_{x \in \mathcal{D}(X^*)} \cos\left(2\pi \left\langle \frac{tx}{2\pi} \right\rangle\right) \mathbb{P}(X^* = x). \tag{3.26}$$

Consider the following bounds on $\cos(x)$ valid for $x \in [0, \pi]$,

$$1 - \frac{x^2}{2} \leq \cos(x) \leq 1 - \frac{2x^2}{\pi^2}.$$

We can use this in combination with (3.26) to deduce that

$$1 - 2\pi^2 \mathbb{E} \left\langle \frac{X^* t}{2\pi} \right\rangle^2 \leq \varphi^*(t) \leq 1 - 8\mathbb{E} \left\langle \frac{X^* t}{2\pi} \right\rangle^2 \tag{3.27}$$

Finally, by definition of X^*, note that $\varphi^*(t) = \varphi(t)\varphi(-t) = |\varphi(t)|^2$. As $|\varphi(t)| \in [0, 1]$ we may deduce that

$$1 - 2\pi^2 H\left(X, \frac{t}{2\pi}\right) \leq |\varphi(t)| \leq 1 - 4H\left(X, \frac{t}{2\pi}\right) \tag{3.28}$$

which may easily be rearranged to give the statement of the lemma. □

For $D \in \mathbb{N}$ and a random variable X we define

$$H_D(X) := \inf_{\frac{1}{4D} \leq d \leq \frac{1}{2D}} H(X, d) \tag{3.29}$$

The following lemma of Mukhin [61] bounds $H(X, d)$ in terms of this new parameter,

Lemma 3.2.6. *For any random variable X, $d \in \mathbb{R}$ and $D \in \mathbb{N}$ with $2D|d| \leq 1$ we have*

$$H(X, d) \geq 4D^2 |d|^2 H_D(X).$$

We may now apply Lemma 3.2.5 to give an explicit upper bound on the integral term in (3.16) as follows. Recall that X_1 is a lattice random variable with step h. By Lemma 3.2.5 and using $\ln(1/x) \geq 1 - x$ for $x > 0$,

$$\int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt \leq \int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} e^{-n(1-|\varphi(t)|)} dt \leq \int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} e^{-4nH(X, \frac{t}{2\pi})} dt. \tag{3.30}$$

Now, note that the upper limit of the integral in (3.30) is π/h. So, as $(\pi/h)/(2\pi) = 1/(2h)$ we may apply Lemma 3.2.6 with $D = h$ and $d = t/2\pi \leq 1/2h$ to deduce that

$$\begin{aligned}
\int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt &\leq \int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} e^{-\frac{4nh^2 H_h(X) t^2}{\pi^2}} dt \leq \int_{\frac{\sigma^2}{4\gamma}}^{\infty} e^{-\frac{4nh^2 H_h(X) t^2}{\pi^2}} dt \\
&= \frac{\pi^{3/2}}{2h(nH_h(X))^{1/2}} \mathbb{P}\left(\mathcal{N}(0,1) > \frac{\sigma^2 h (nH_h(X))^{1/2}}{\sqrt{2}\pi\gamma}\right) \\
&\leq \frac{\pi^{3/2}}{2h(nH_h(X))^{1/2}} e^{-\frac{\sigma^4 h^2 n H_h(X)}{4\pi^2\gamma^2}},
\end{aligned} \tag{3.31}$$

where the final inequality follows by the Chernoff's bound. This allows us to deduce, once again using the inequality $e^{-x} < x^{-1/2}$, that this integral is bounded above as

$$\int_{\frac{\sigma^2}{4\gamma}}^{\frac{\pi}{h}} |\varphi(t)|^n dt \leq \frac{\pi^{5/2}\gamma}{h^2 \sigma^2 n H_h(X)}.$$

This allows us to state the following corollary to Theorem 3.2.1.

Corollary 3.2.7. Let $X_1, X_2, \ldots, X_n$ be independent and identically distributed random variables taking values on $\mathcal{L}(v_0, h)$. Define $S_n = \sum_{i=1}^n X_i$. Suppose that $\mu = \mathbb{E}(X_1) = 0$, $\sigma^2 = \mathrm{Var}(X_1)$ and $\gamma = \mathbb{E}|X_1|^3$. Then,

$$\sup_{w \in \mathcal{L}(nv_0,h)} \left| \mathbb{P}(S_n = w) - \frac{h}{\sqrt{2\pi n\sigma^2}} \exp\left(-\frac{w^2}{2n\sigma^2}\right) \right| \leq \frac{32h\gamma}{\sigma^4 n} + \frac{6\gamma}{h\sigma^2 n H_h(X_1)}. \tag{3.32}$$

To give an explicit upper bound on the error probability, we need to deduce that $H_h(X_1)$ is bounded from below. For $\mathbf{x} = (x_1, x_2, \ldots, x_k) \in \mathbb{Z}^k$, define

$$w(\mathbf{x}) := \max_{i \neq \ell \neq j} \frac{|x_i - x_\ell|}{\gcd(|x_i - x_\ell|, |x_j - x_\ell|)}.$$

Then the fact that $H_h(X_1)$ is bounded from below is implied by the following lemma,

Lemma 3.2.8. *Let X be an integer valued random variable supported on a lattice of step h and with atoms $x_1, \ldots, x_k$ not all contained in a non-trivial arithmetic progression of the lattice. Then there exists an absolute constant $C > 0$ such that*

$$H_h(X) \geq C \cdot \frac{\min_{i \in [k]} \mathbb{P}(X = x_i)}{k(w(\mathbf{x})h)^2}$$

Proof. For $d \in \mathbb{R}$, consider

$$D(X, d) := \inf_{\alpha \in \mathbb{R}} \mathbb{E}\langle (X - \alpha)d \rangle^2. \tag{3.33}$$

By [61, Lemma 1], we have that $D(X, d) \leq H(X, d) \leq 4D(X, d)$. Therefore,

$$H_h(X) \geq \min_{1/4h \leq d \leq 1/2h} D(X, d)$$

For all $\mathbf{x} \in \mathbb{Z}^k$ and $\beta, d \in \mathbb{R}$, define

$$S(\mathbf{x}, \beta, d) := \sum_{i=1}^k \langle \beta + x_i d \rangle \tag{3.34}$$

Cauchy-Schwartz's inequality implies that

$$\begin{aligned} D(X,d) &:= \inf_{\alpha\in\mathbb{R}} \mathbb{E}\langle (X-\alpha)d\rangle^2 \\ &= \inf_{\beta\in\mathbb{R}} \sum_{i=1}^{k} \langle \beta + x_i d\rangle^2 \mathbb{P}(X = x_i) \\ &\geq \min_{i\in[k]} \mathbb{P}(X = x_i) \inf_{\beta\in\mathbb{R}} \sum_{i=1}^{k} \langle \beta + x_i d\rangle^2 \\ &\geq \frac{\min_{i\in[k]} \mathbb{P}(X = x_i)}{k} \inf_{\beta\in\mathbb{R}} S(\mathbf{x},\beta,d)^2. \end{aligned} \tag{3.35}$$

It thus suffices to bound the infimum of S when $\mathbf{x}$ and d are fixed. The derivative of S with respect to β satisfies the following properties:

(i) it is well defined for all β such that $\langle \beta + x_i d\rangle \notin \{0, 1/2\}$ for all $i \in [k]$;

(ii) it is constant between any two consecutive values at which the derivative is undefined;

(iii) it takes integer values in $\{-k, \ldots, k\}$ anywhere where it is defined.

Thus, the minimum of S is attained at β_0, for which the derivative is not defined. By relabelling the x_i, we may assume that $\langle \beta_0 + x_k d\rangle \in \{0, 1/2\}$. If $\langle \beta_0 + x_k d\rangle = 1/2$, plugging it in (3.35) we would get the desired bound done. So we may assume that $\langle \beta_0 + x_k d\rangle = 0$, and, in fact, we can choose $\beta_0 = -x_k d$. For $i \in [k-1]$, define $y_i := x_i - x_k$. As the x_i are not all contained in a non-trivial arithmetic progression then $\mathrm{hcf}(y_1, y_2, \ldots, y_{k-1}) = h$. By a simple extension of Bézout's Lemma there exist $\lambda_i \in \mathbb{Z}$ with $|\lambda_i| \leq w(\mathbf{x})h$ for all $i \in [k-1]$ and $\lambda_1 y_1 + \lambda_2 y_2 + \ldots + \lambda_{k-1} y_{k-1} = h$. Now, using the identities $\langle t\beta\rangle \leq |t|\langle \beta\rangle$ for any $t \in \mathbb{Z}$ and $\langle \beta_1 + \beta_2\rangle \leq \langle \beta_1\rangle + \langle \beta_2\rangle$, we obtain

$$\inf_{\beta\in\mathbb{R}} S(\mathbf{x},\beta,d) = \sum_{i=1}^{k} \langle \beta_0 + x_i d\rangle = \sum_{i=1}^{k-1} \langle y_i d\rangle \geq \sum_{i=1}^{k-1} \frac{\langle \lambda_i y_i d\rangle}{|\lambda_i|} \geq \frac{\left\langle \sum_{i=1}^{k-1} \lambda_i y_i d \right\rangle}{w(\mathbf{x})h} = \frac{\langle hd\rangle}{w(\mathbf{x})h} \tag{3.36}$$

Observing that $\langle hd \rangle = hd$ for all $d \leq 1/2h$, we obtain,

$$H_h(X) \geq \min_{1/4h \leq d \leq 1/2h} D(X,d) \geq \min_{1/4h \leq d \leq 1/2h} \frac{\min_{i \in [k]} \mathbb{P}(X = x_i)}{k} \left(\frac{\langle hd \rangle}{w(\mathbf{x})h} \right)^2 = \frac{\min_{i \in [k]} \mathbb{P}(X = x_i)}{16k(w(\mathbf{x})h)^2}.$$

□

3.3 Barely subcritical regime for the configuration model

3.3.1 Exploration process

In this section we introduce a process that given a vertex $v \in [n]$ explores $\mathbb{CM}_n$ starting by the component containing v. We set a total order of the half edges as follows. For every vertex v, consider an arbitrary order of its d_v half-edges. Then, the half edges are ordered, first by its corresponding vertex (using the total order on $[n]$) and then by the order given within the half-edges incident to a vertex.

We will denote by $\mathcal{F}_t$ the history of the process at time t. With a slight abuse of notation, we will assume that $\mathcal{F}_t$ is the subgraph formed by the partial matching at time t. Note that the order of the pairings is determined by the knowledge of the matching. The main random variable we would like to track is $X_t = X_t(v)$, defined as the number of unmatched half-edges incident to $V(\mathcal{F}_t)$ when the process started at v. Note that if $X_t = 0$, there are no unpaired half-edges and thus $\mathcal{F}_t$ is a union of components of $\mathbb{CM}_n$ containing the component of v.

The *exploration process of* $\mathbb{CM}_n$ *starting at* $v \in [n]$ is defined as follows:

1) Let $\mathcal{F}_0$ be the single-vertex graph on $\{v\}$ and $X_0 = d_v$.

2) While $V(\mathcal{F}_t) \neq [n]$,

   2a) If $X_t = 0$, choose a uniformly unmatched half-edge and let u be the vertex incident to it. Let $\mathcal{F}_{t+1}$ be constructed from $\mathcal{F}_t$ by adding $\{u\}$ as an isolated vertex, and let $X_{t+1} = d_u$.

2b) Otherwise, choose the smallest unmatched half-edge e incident to $V(\mathcal{F}_t)$ and pair it with a half-edge f chosen uniformly at random from all the unmatched ones. Let u be the vertex incident to f.

i) If $u \notin V(\mathcal{F}_t)$, let $\mathcal{F}_{t+1}$ be constructed from $\mathcal{F}_t$ by adding vertex u and edge ef and let $X_{t+1} = X_t + d_u - 2$.

ii) Otherwise, let $\mathcal{F}_{t+1}$ constructed from $\mathcal{F}_t$ by adding edge ef and let $X_{t+1} = X_t - 2$.

Note that X_t is measurable with respect to $\mathcal{F}_t$. We define the following parameters:

$$\begin{aligned} \eta_{t+1} &:= X_{t+1} - X_t \ , \\ M_t &:= X_t + \sum_{u \notin V(\mathcal{F}_t)} d_u \ , \\ Q_t &= \frac{1}{M_t - 1} \sum_{u \notin V(\mathcal{F}_t)} d_u(d_u - 2) \ , \\ R_t &= \frac{1}{M_t - 1} \sum_{u \notin V(\mathcal{F}_t)} d_u(d_u - 2)^2 \ . \end{aligned} \tag{3.37}$$

It is straightforward to check that if $X_t > 0$, then

$$\mathbb{E}\left[\eta_{t+1} \mid \mathcal{F}_t\right] = Q_t \quad \text{and} \quad \mathbb{E}\left[(\eta_{t+1})^2 \mid \mathcal{F}_t\right] = R_t \ , \tag{3.38}$$

and if $X_t = 0$, then

$$\mathbb{E}\left[\eta_{t+1} \mid \mathcal{F}_t\right] = \frac{1}{M_t} \sum_{u \notin V(\mathcal{F}_t)} d_u^2 \quad \text{and} \quad \mathbb{E}\left[(\eta_{t+1})^2 \mid \mathcal{F}_t\right] \geq \frac{R_t}{2} \ , \tag{3.39}$$

although we will never study the process for t such that $X_t = 0$.

3.3.2 Stochastic domination and random sums

Recall the definition of $T = T_n$ given in Equation (3.6).

Define the distribution β as follows: for every $\ell \in L := \{-1, 0, 1, \ldots, n-3\}$,

$$\mathbb{P}(\beta = \ell) := \begin{cases} \frac{m}{m-2T}\,\mathbb{P}(\eta = -1) - \frac{2T}{m-2T} & \text{if } \ell = -1\ , \\ \frac{m}{m-2T}\,\mathbb{P}(\eta = \ell) & \text{if } \ell \geq 0\ . \end{cases} \tag{3.40}$$

Let $\varphi_\beta(\theta)$, θ_0^β, Q_β, R_β and T_β be defined as in Equations (3.1), (3.2), (3.5) and (3.6) replacing η by β. By the choice of β all main parameters are asymptotically equal to the original ones.

Lemma 3.3.1. *For every $k \geq 0$, we have $\varphi_\beta^{(k)}(\theta) = (1 + o(\theta_0))\varphi^{(k)}(\theta) + o(\theta_0)$. Moreover, $T_\beta = (1 + o(1))T$.*

Proof. The first part of the lemma follows directly from

$$\begin{aligned} \varphi_\beta^{(k)}(\theta) = \mathbb{E}(\beta^k e^{\theta\beta}) &= \frac{m}{m-2T}\mathbb{E}(\eta^k e^{\theta\eta}) - \frac{2T}{m-2T}(-1)^k e^{-\theta} && (3.41) \\ &= (1 + O(T/m))\varphi^{(k)}(\theta) + O(T/m) \\ &= (1 + o(\theta_0))\varphi^{(k)}(\theta) + o(\theta_0). && (3.42) \end{aligned}$$

where in the last line we used the hypothesis $T = o(m\theta_0)$ in Theorem 3.1.5.

For the second part, we split into two cases. If $\Delta|Q| = o(R)$, we are in the setting of Theorem 3.1.3. In such case

$$\log\Big(\varphi_\beta(\theta_0^\beta)\Big)^{-1} \sim \frac{Q_\beta^2}{2R_\beta} \sim \frac{Q^2}{2R} \sim \log(\varphi(\theta_0))^{-1},$$

(see Section 3.3.4 for the first and third equivalences) and the result follows from the first part of the lemma.

Otherwise $R = O(\Delta|Q|)$. As R is bounded away from zero by Assumption 3.1.1 and $\Delta_n \leq n^{1/6}$, it follows that $|Q|$ is of order at least $n^{-1/6}$. Again all we need to show is that $\log\Big(\varphi_\beta(\theta_0^\beta)\Big)^{-1} = (1 + o(1))\log(\varphi(\theta_0))^{-1}$ and then the rest will follow by the first part of the lemma. We do this by bounding $\varphi(\theta_0) - \varphi_\beta(\theta_0^\beta)$.

Using $e^x \geq 1 + x$, we have

$$0 = \mathbb{E}[\eta e^{\theta_0 \eta}] \geq \mathbb{E}[\eta(1 + \theta_0 \eta)] = Q + \theta_0 R$$

and thus

$$0 < \theta_0 \leq \frac{|Q|}{R} = o(1). \tag{3.43}$$

By construction, β stochastically dominates η and it follows that $\varphi_\beta^{(k)}(\theta) \geq \varphi^{(k)}(\theta)$ for all $k \geq 0$ and $\theta \geq 0$. In particular,

$$0 < \theta_0^\beta \leq \theta_0 \tag{3.44}$$

Combining (3.41) for $k = 0$, (3.43) and (3.44),

$$\varphi_\beta(\theta_0^\beta) - \varphi(\theta_0^\beta) = \frac{2T}{m - 2T}(\varphi(\theta_0^\beta) - e^{-\theta_0^\beta}) \leq \frac{2T\theta_0}{m}(1 + o(1)) = O\left(\frac{T|Q|}{mR}\right). \tag{3.45}$$

As φ_β'' is an increasing function with $\varphi_\beta''(0) = (1+o(1))R$ and $\varphi_\beta'(\theta_0^\beta) = 0$, the fundamental theorem of calculus implies

$$(\theta_0 - \theta_0^\beta)R \leq (1 + o(1)) \int_{\theta_0^\beta}^{\theta_0} \varphi_\beta''(t)dt = (1 + o(1))\varphi_\beta'(\theta_0) = O\left(\frac{T}{m}\right). \tag{3.46}$$

where the last equality follows from (3.42).

We have $|\varphi'(t)| \leq |Q|$ for all $t \in [0, \theta_0]$; indeed, φ' is increasing with $\varphi'(0) = Q$ and $\varphi'(\theta_0) = 0$, Similarly as before, using (3.46) we conclude that

$$\varphi(\theta_0) - \varphi(\theta_0^\beta) = \int_{\theta_0^\beta}^{\theta_0} \varphi'(t)dt \leq (\theta_0 - \theta_0^\beta)|Q| = O\left(\frac{T|Q|}{mR}\right). \tag{3.47}$$

Combining (3.45) and (3.47),

$$\varphi_\beta(\theta_0^\beta) - \varphi(\theta_0) = O\left(\frac{T|Q|}{mR}\right). \tag{3.48}$$

Recall that $\varphi'(0) = Q < 0$. Using the inequality $e^x \leq 1 + 2x$ for $x \in [0,1]$, we have

$$\varphi'\left(\frac{|Q|}{5\Delta}\right) \leq \mathbb{E}\left[\eta\left(1 + \frac{2|Q|\eta}{5\Delta}\right)\right] \leq Q + \frac{2|Q|R}{5\Delta} < 0$$

where the final inequality holds because $Q < 0$ implies $R < 2\Delta$. It follows that $\theta_0 \geq \frac{|Q|}{5\Delta}$.

By using the fact that $\varphi'(\theta) < 0$ for all $\theta \in [0, \theta_0)$ and Taylor expansion of $\varphi(\theta)$ around $\theta = 0$, we obtain

$$\begin{aligned}
\varphi(\theta_0) \leq \varphi\left(\frac{|Q|}{5\Delta}\right) &\leq 1 - \frac{Q^2}{5\Delta} + \sum_{k\geq 2} \frac{\mathbb{E}[\eta^k]}{k!} \cdot \left(\frac{|Q|}{5\Delta}\right)^k \\
&\leq 1 - \frac{Q^2}{5\Delta} + \frac{2Q^2}{25\Delta} \sum_{\ell \geq 0} \left(\frac{|Q|}{5}\right)^\ell \\
&\leq 1 - \frac{Q^2}{10\Delta}.
\end{aligned} \tag{3.49}$$

where in the third inequality we used that $\mathbb{E}[\eta^k] \leq 2\Delta^{k-1}$ for all $k \in \mathbb{N}$, since $Q < 0$.

Using the bound (3.49) in the definition of T gives the simple upper bound $T \leq \frac{10\Delta}{Q^2}\log(n)$. By our bounds on Δ and Q, $T^2|Q| = O(n^{-5/6}\log n) = o(mR)$. Thus, substituting this into (3.48), we have

$$\varphi_\beta(\theta_0^\beta) - \varphi(\theta_0) = o\left(\frac{1}{T}\right). \tag{3.50}$$

By definition, $\log(\varphi(\theta_0))^{-1} \geq 1/T$. Therefore,

$$\varphi(\theta_0) \leq e^{1/T} = 1 - \frac{1 + o(1)}{T} \tag{3.51}$$

Combining (3.50) and (3.51),

$$\log\Big(\varphi_\beta(\theta_0^\beta)\Big)^{-1} \sim 1-\varphi_\beta(\theta_0^\beta) \sim 1-\varphi(\theta_0) \sim \log(\varphi(\theta_0))^{-1},$$

concluding the proof of the lemma.

□

Let $(\beta_t)_{t\geq 1}$ be a sequence of iid copies of β. For $s \in N$, define the stochastic process $W_t = W_t^s$ by $W_0 = s$ and for $t \geq 0$

$$W_{t+1} = W_t + \beta_t = s + \sum_{i=1}^{t} \beta_i \ . \tag{3.52}$$

Define the stopping time

$$\tau_W^s := \inf\{t : W_t^s = 0\} \ .$$

Let h be the largest possible common difference of a progression upon which the limiting degree sequence D is supported, that is

$$h := \max\{j : \exists k \text{ s.t. } \mathbb{P}(D \in \mathcal{L}(k,j)) = 1\}$$

Lemma 3.3.2. *For every $t \geq T_\beta$ and $s = s(n)$ we have*

$$\mathbb{P}\left(\tau_W^s = t\right) \leq 2h \cdot s e^{\theta_0^\beta s} \left(\varphi_\beta''(\theta_0^\beta)\right)^{-1/2} \frac{(\varphi_\beta(\theta_0^\beta))^t}{t^{3/2}} \ . \tag{3.53}$$

Moreover, for every $\epsilon > 0$ we have that

$$\mathbb{P}\left(\tau_W^s \geq (1+\epsilon)T_\beta\right) = o\left(\frac{s e^{\theta_0^\beta s}}{\mathbb{E}\left[D_n e^{\theta_0^\beta D_n}\right]} \cdot \frac{T_\beta}{n}\right) \ . \tag{3.54}$$

Proof. The dependence on s is implicit in all the notation below. Define the following sequences

$$\begin{aligned} I_t &= \{\mathbf{b} = (b_1, \ldots, b_t) \in L^t : \, s + b_1 + \cdots + b_t = 0\} \\ \hat{I}_t &= \{\mathbf{b} = (b_1, \ldots, b_t) \in I_t : \, s + b_1 + \cdots + b_i > 0, \forall i \in [t-1]\} \end{aligned} \tag{3.55}$$

We can write

$$\mathbb{P}\left(W_t = 0\right) = \sum_{\mathbf{b} \in I_t} \prod_{i=1}^{t} \mathbb{P}\left(\beta_i = b_i\right) \quad \text{and} \quad \mathbb{P}\left(\tau_W^s = t\right) = \sum_{\mathbf{b} \in \hat{I}_t} \prod_{i=1}^{t} \mathbb{P}\left(\beta_i = b_i\right) \tag{3.56}$$

A variant of Spitzer's lemma [63, Lemma 9] implies that

$$\mathbb{P}\left(\tau_W^s = t\right) \leq \frac{s}{t} \mathbb{P}\left(W_t = 0\right) \tag{3.57}$$

We will use exponential tilting to bound the probability that $W_t = 0$, as in [63, 72]. Consider the probability distribution β_θ defined for $\ell \in L$ by

$$\mathbb{P}\left(\beta_\theta = \ell\right) = \frac{e^{\theta \ell} \mathbb{P}\left(\beta = \ell\right)}{\varphi_\beta(\theta)} \, . \tag{3.58}$$

Let $(\beta_{\theta,t})_{t \geq 1}$ be a sequence of iid copies of β_θ. Define the stochastic process $W_{\theta,t}$ by $W_{\theta,0} = s$ and for $t \geq 0$

$$W_{\theta,t+1} = s + \sum_{i=1}^{t} \beta_{\theta,i} \, . \tag{3.59}$$

Algebraic manipulations give

$$\mathbb{P}\left(W_t = 0\right) = (\varphi_\beta(\theta))^t e^{\theta s} \mathbb{P}\left(W_{\theta,t} = 0\right) \, . \tag{3.60}$$

By definition of θ_0^β, $\mathbb{E}\left[\beta_{\theta_0^\beta}\right] = \mathbb{E}\left[\beta e^{\beta \theta_0^\beta}\right] = 0$. We may write $W_{\theta_0^\beta,t} = s + S_t$, where

$S_t = \sum_{i=1}^t Y_i$ and $(Y_i)_{i\in[t]}$ is a collection of iid copies of $\beta_{\theta_0^\beta}$. In particular, we have

$$\begin{aligned}
\mu &= \mathbb{E}\left[Y_1\right] = \varphi'_\beta(\theta_0^\beta) = 0\ , \\
\sigma^2 &= \mathbb{E}\left[Y_1^2\right] = \frac{\varphi''_\beta(\theta_0^\beta)}{\varphi_\beta(\theta_0^\beta)}\ , \\
\gamma &= \mathbb{E}\left[|Y_1|^3\right] \le 2 + \mathbb{E}\left[Y_1^3\right] = \frac{2\varphi_\beta(\theta_0^\beta) + \varphi'''_\beta(\theta_0^\beta)}{\varphi_\beta(\theta_0^\beta)}\ .
\end{aligned} \tag{3.61}$$

where in the inequality, we have used that $Y_1 \ge -1$.

We will apply Corollary 3.2.7 and show that the error term is negligible with respect to the Gaussian probability. Recall that h is the step of the limiting distribution D, which by Assumption 3.1.1 is also the step of the distribution of Y_1. Since h and $H_h(Y_1)$ (as defined in (3.29)) are constants, the order of the first error term in (3.32) is at most the order of the second one, and it suffices to bound the latter. Assumption 3.1.1 implies that $\sigma^2 \ge \mathbb{P}(\hat{D}_n \ne 2) > 0$ for large n, and that $\gamma = O(\Delta^{1/2})$. Therefore, for any $t \ge T_\beta$

$$\frac{\gamma}{\sigma^2 t} = O\left(\frac{1}{\sqrt{\sigma^2 t}} \cdot \sqrt{\frac{\Delta}{T_\beta}}\right) = o\left(\frac{1}{\sqrt{\sigma^2 t}}\right).$$

where we used that $\Delta = o(T)$ and $T \sim T_\beta$ by Lemma 3.3.1.

Since $\mathbb{P}\left(Y_1 = -1\right) > 0$, we may choose $v_0 = -1$. Thus, for sufficiently large n, we conclude that for any $w \in \mathcal{L}(-t, h)$,

$$\mathbb{P}(S_t = w) \le \frac{2h}{\sqrt{2\pi t \sigma^2}} \tag{3.62}$$

We can now use (3.62) with $w = -s$ to obtain

$$\mathbb{P}\left(W_{\theta_0,t} = 0\right) = \mathbb{P}\left(S_t = -s\right) \le \frac{2h}{\sqrt{2\pi t}} \left(\frac{\varphi_\beta(\theta_0^\beta)}{\varphi''_\beta(\theta_0^\beta)}\right)^{1/2} \tag{3.63}$$

Let us show that $\varphi_\beta(\theta_0^\beta)$ is close to 1. On the one hand, we will use the inequality $e^x \le 1 + xe^x$ for all $x \in \mathbb{R}$, with equality if and only if $x = 0$. Since $\mathbb{P}\left(\beta = 0\right) \ne 1$, by the

choice of θ_0^β

$$\varphi_\beta(\theta_0^\beta) = \mathbb{E}\left[e^{\theta_0^\beta \beta}\right] < 1 + \theta_0^\beta \mathbb{E}\left[\beta e^{\theta_0^\beta \beta}\right] = 1 \tag{3.64}$$

On the other hand, using $e^x \geq 1 + x$ for $x \in \mathbb{R}$, that θ_0^β is bounded on n and $\mathbb{E}[\beta] = o(1)$, we obtain

$$\varphi_\beta(\theta_0^\beta) \geq 1 + \theta_0^\beta \mathbb{E}[\beta] = 1 + o(1) \ . \tag{3.65}$$

Thus, we use the asymptotic equivalence $\log\Big(\varphi_\beta(\theta_0^\beta)\Big)^{-1} \sim 1 - \varphi_\beta(\theta_0^\beta)$.

Combining Equations (3.57), (3.60), (3.63) and (3.64), we obtain

$$\mathbb{P}\left(\tau_W^s = t\right) \leq 2h \cdot s e^{\theta_0^\beta s} \left(\varphi_\beta''(\theta_0^\beta)\right)^{-1/2} \frac{(\varphi_\beta(\theta_0^\beta))^t}{t^{3/2}} \ .$$

proving the first part of the lemma.

For the second statement of the lemma, it suffices to prove it for small enough ϵ, so we may assume $\varepsilon \in (0,1)$. Observe that $\mathbb{P}\left(\tau_W^s = t\right) \neq 0$ implies that $s = hk - t$ for some $k \in \mathbb{Z}$. Since $v_0 = -1$ and h are coprime, there are at most $\lceil T/h \rceil$ values $t \in [T]$ such that $\mathbb{P}\left(\tau_W^s = t\right) \neq 0$.

As our bound on $\mathbb{P}\left(\tau_W^s = t\right)$ is decreasing on t, using Equation (3.64) it follows that

$$\begin{aligned}
\mathbb{P}\left(\tau_W^s \geq (1+\epsilon)T_\beta\right) &= \sum_{\substack{t \geq (1+\epsilon)T_\beta \\ t \in \mathcal{L}(-1,h)}} \mathbb{P}\left(\tau_W^s = t\right) \leq 2 s e^{\theta_0^\beta s} \left(\varphi_\beta''(\theta_0^\beta)\right)^{-1/2} \frac{(\varphi_\beta(\theta_0^\beta))^{(1+\epsilon)T_\beta}}{T_\beta^{3/2}} \sum_{\ell \geq 0} (\varphi_\beta(\theta_0^\beta))^\ell \\
&= 2 s e^{\theta_0^\beta s} \left(\varphi_\beta''(\theta_0^\beta)\right)^{-1/2} \frac{(\varphi_\beta(\theta_0^\beta))^{(1+\epsilon)T_\beta}}{T_\beta^{3/2}(1 - \varphi_\beta(\theta_0^\beta))} = o\left(\frac{s e^{\theta_0^\beta s}}{\mathbb{E}\left[D_n e^{\theta_0^\beta D_n}\right]} \cdot \frac{T_\beta}{n}\right) \ ,
\end{aligned}$$

where in the last equality we used that $T_\beta \sim \frac{1}{\log \varphi_\beta^{-1}(\theta_0^\beta)} \log\left(T_\beta^{-3/2} (\varphi_\beta''(\theta_0^\beta))^{-1/2} \mathbb{E}\left[D_n e^{\theta_0^\beta D_n}\right] n\right)$.

□

3.3.3 Proof of Theorem 3.1.5

Fix $\epsilon > 0$ sufficiently small. Define the stopping time $\tau_X(v)$ as the number of edges in the component of v, denoted by $\mathcal{C}(v)$. That is,

$$\tau_X(v) := \inf\{t : X_t(v) = 0\} .$$

Note that for every $t \leq 2T \wedge \tau_X(v)$, the distribution β stochastically dominates η_t. Thus, $X_t(v)$ is stochastically dominated by $W_t^{d_v}$.

Let $\delta = \epsilon/3$. It follows from Lemmas 3.3.1 and 3.3.2 that

$$\mathbb{P}\left(\tau_X(v) \geq (1+2\delta)T\right) \leq \mathbb{P}\left(\tau_X \geq (1+\delta)T_\beta\right) \leq \mathbb{P}\left(\tau_W^s \geq (1+\delta)T_\beta\right) = o\left(\frac{d_v e^{\theta_0^\beta d_v}}{\mathbb{E}\left[D_n e^{\theta_0^\beta D_n}\right]} \cdot \frac{T}{n}\right) . \tag{3.66}$$

Let Z be the number of components of order at least $(1+\epsilon)T$. For any $\epsilon > 0$, we can write

$$Z = \sum_{\mathcal{C}} \mathbb{1}|\mathcal{C}| \geq (1+\epsilon)T = \sum_{v\in[n]} \frac{\mathbb{1}|\mathcal{C}(v)| \geq (1+\epsilon)T}{|\mathcal{C}(v)|} \leq \frac{1}{T}\sum_{v\in[n]} \mathbb{1}|\mathcal{C}(v)| \geq (1+\epsilon)T , \tag{3.67}$$

where the first sum is over the connected components of $\mathbb{CM}_n$.

Since $\mathcal{C}(v)$ is a connected subgraph, it has at least $|\mathcal{C}(v)| - 1$ edges. Thus, the probability of $|\mathcal{C}(v)| \geq k$ is bounded from above by the probability $\tau_X(v) \geq k-1$. Using Equation (3.66) we obtain

$$\mathbb{E}\left[Z\right] \leq \frac{1}{T}\sum_{v\in[n]} \mathbb{P}\left(\tau_X(v) \geq (1+\epsilon)T - 1\right) = o\left(\frac{1}{\mathbb{E}\left[D_n e^{\theta_0^\beta D_n}\right] n}\right) \sum_{v\in[n]} d_v e^{\theta_0^\beta d_v} = o(1) .$$

Theorem 3.1.5 follows by Markov's inequality on Z.

3.3.4 Proof of Theorem 3.1.3

Recall that $\Delta|Q| = o(R)$ and that $\varphi(\theta)$ is the moment generating function of η. Thus, $\varphi(0) = 1$, $\varphi'(0) = Q$, $\varphi''(0) = R$ and $\varphi^{(k)}(0) \leq \Delta^{k-3}R$ for all $k \geq 3$. This implies that the radius of convergence of φ (and so of any of its derivatives) is at least $2|Q|/R$. So, for any θ with $|\theta| < 2|Q|/R$, we have

$$\varphi'(\theta) = \varphi'(0) + \theta\varphi''(0) + O(\theta^2\varphi'''(0)) = Q + \theta R + o(Q) \ .$$

By the choice of θ_0, we have $\varphi'(\theta_0) = 0$ and $\theta_0 \sim |Q|/R$.

We can also write

$$\varphi(\theta_0) = \varphi(0) + \theta_0\varphi'(0) + \frac{\theta_0^2\varphi''(0)}{2} + o(1) \sim 1 - \frac{Q^2}{2R} \ .$$

and

$$\log(\varphi(\theta_0))^{-1} \sim \frac{Q^2}{2R} \ .$$

Similar arguments give that $\varphi''(\theta_0) \sim R$.

Finally, observe that for any $\delta > 0$,

$$\mathbb{E}\left[De^{D\theta_0}\right] \leq \mathbb{E}\left[De^{(1+\delta)\Delta Q/R}\right] = \mathbb{E}\left[De^{o(1)}\right] = O(1) \ . \tag{3.68}$$

Using all previous estimations, we can write

$$T \sim \frac{2R}{Q^2}\log\left(\frac{|Q|^3 n}{R^2}\right) \ .$$

It is straightforward to check that, in this case, the condition $\theta_0 m \geq \omega(n)T_n$ is equivalent to $Q \leq -\omega(n)n^{-1/3}R^{2/3}$.

Note that the condition $\Delta_n \leq n^{1/6}$ is only required in Lemma 3.3.1 in the case

$R = O(\Delta|Q|)$. Thus, the desired result follows from Theorem 3.1.5 without further restrictions on the degree sequence.

3.4 Subcritical regime for the uniform model

3.4.1 Exploration process

We will use the exploration process described in [45] that, given $V_0 \subset [n]$, reveals the components of $\mathbb{G}_n$ one by one starting with the components containing V_0.

We first describe the exploration process on fixed graphs where each vertex has an order in its adjacency list. Precisely, an *input* is a pair (G, Π), with G a graph on $[n]$ and $\Pi = (\pi_v)_{v\in[n]}$ a collection of permutations where π_v has length d_v and induces a natural order on the edges incident to v. The process constructs a sequence of sets $V_0 \subset V_1 \subset \dots$ such that at time t all the edges in $G[V_t]$ have been revealed. Similarly as before, we define $X_t = X_t(v) = |E(V_t, [n] \setminus V_t)|$ to be the number of edges between the explored and unexplored parts. If $X_t = 0$, V_t is a set of vertices forming a union of components, including the ones intersecting V_0. We also define $E(A, B)$ to be the set of edges between sets A and B, $M_t = \sum_{w\in[n]\setminus V_t} d_w$, and we let L_t be the number of vertices of degree 1 in $[n] \setminus V_t$.

The *exploration process of* (G, Π) *starting at* $V_0 \subset [n]$ is defined as follows:

1) Let $X_0 = |E(V_0, [n] \setminus V_0)|$.

2) While $V_t \neq [n]$,

   2a) If $X_t = 0$, choose a vertex u in $[n] \setminus V_t$ according to the degree distribution and let $w_{t+1} = u$, i.e. $\mathbb{P}(w_{t+1} = u) = \frac{d_u}{M_t}$. Let $V_{t+1} = V_t \cup \{w_{t+1}\}$ and $X_{t+1} = d_{w_{t+1}}$.

   2b) Otherwise, choose v_{t+1} the smallest vertex incident to at least one edge in $[n]\setminus V_t$ and let e_{t+1} be the smallest edge in $E(v_{t+1}, [n] \setminus V_t)$. Let w_{t+1} be the endpoint

of e_{t+1} in $[n] \setminus V_t$. Expose all edges in $E(w_{t+1}, V_t)$. Let $V_{t+1} = V_t \cup \{w_{t+1}\}$ and $X_{t+1} = X_t - 1 + d_{w_{t+1}} - |E(w_{t+1}, V_t)|$.

There are two main differences between this exploration process and the one defined in Section 3.3.1: we explore vertex by vertex instead of edge by edge, and we start from a set instead of a single vertex.

We will run the exploration process on an input (G, Π) chosen uniformly at random from all the inputs where G is a graph on $[n]$ with degree sequence $\mathbf{d}_n$. This is equivalent to sampling $G \sim \mathbb{G}_n$ and, independently, letting $\Pi = \Pi(\mathbf{d}_n)$ be a collection of uniformly and independent permutations of lengths $(d_v)_{v\in[n]}$. We will use the *principle of deferred decisions* exposing the restriction of $\pi_{v_{t+1}}$ onto $E(v_{t+1}, [n] \setminus V_t)$ at time t. Let $(\mathcal{F}_t)_{t\geq 0}$ be the filtration of the space of inputs given by the history of the process just after exposing the order on $E(v_{t+1}, [n] \setminus V_t)$. The random objects X_t, V_t, M_t, L_t, v_{t+1} and e_{t+1} are $\mathcal{F}_t$-measurable, while w_{t+1} is $\mathcal{F}_{t+1}$-measurable. We will use $\mathbb{P}_t(\cdot) := \mathbb{P}(\cdot \mid \mathcal{F}_t)$ and $\mathbb{E}_t[\cdot] := \mathbb{E}[\cdot \mid \mathcal{F}_t]$ to denote respectively the probability and expected value conditioned to $\mathcal{F}_t$.

3.4.2 Deterministic properties of the process

First of all, we may assume that $m_0 = o(n)$, as otherwise since $|Q_0| \leq 1$, there is nothing to prove. This implies that, $m \leq 3n$ as $\sum_{w\in[n]\setminus S_*} d_w \leq 2n$.

Define

$$T := \frac{m_0}{|Q_0|} \geq \frac{\Delta|Q_0| + R}{Q_0^2} \log \lambda \tag{3.69}$$

and

$$\lambda := \frac{nQ_0^2}{\Delta|Q_0| + R} \geq \frac{\omega(n)}{|Q_0|} \to \infty\,. \tag{3.70}$$

The last condition imposed implies that $T = o(|Q|m)$, we will use this bound repeatedly

during the proof.

Let S be the set that certifies (m_0, Q_0)-subcriticality. We can always assume that such set is formed by vertices of largest degrees. Note that our condition on m_0 implies that $\Delta = o(m_0)$. So, we may also assume that

$$\sum_{v \in S} d_v \geq \frac{m_0}{2} , \tag{3.71}$$

as increasing the set S can only decrease Q_0.

Throughout the proof, we will assume that

$$\Delta' := \max_{w \in [n] \setminus S} d_w \geq 2 . \tag{3.72}$$

Otherwise there are at most m_0 vertices of degree at least 2 and any component has order at most $O(m_0)$ and we are done.

Lemma 3.4.1. *Let $v \in [n]$ and set $V_0 = S \cup \{v\}$. We have:*

1. $\sum_{u \in V_0} d_u \leq 2|Q_0|T$;
2. $\Delta' = o\left(\frac{n_1}{\Delta}\right)$.

Moreover, for every $t = O(T)$ we have:

4. $n_1(t) \geq n_1/2$;
5. $M_t \geq m/3$.

Proof. For Item 1, just observe that $d_v \leq \Delta \leq m_0 = |Q_0|T$.

For Item 2, since $\Delta' \geq 2$, $m_0 \geq 3m_*$ and by Equations (3.9), (3.10) and (3.71), we have

$$0 \geq \sum_{v \in [n] \setminus S_*} d_v(d_v - 2) \geq -n_1 + (\Delta' - 2) \sum_{w \in S \setminus S_*} d_w \geq -n_1 + (\Delta' - 2)(m_0/2 - m_*) ,$$

From here it follows that $\Delta' = O(n_1/m_0) = o(n_1/\Delta)$.

For Item 4, observe that $n_1 \geq |Q_0|m$ and $T = o(|Q_0|m)$. So $n_1(t) \geq n_1 - 1 - t \geq n_1/2$.

For Item 5, by Equation (3.9) we have

$$0 \geq Qm \geq -n_1 + \sum_{\substack{v \in [n] \setminus S_* \\ d_v \geq 3}} d_v = m - m_* - 2(n_1 + n_2) \ .$$

Counting only the contribution of vertices of degree 1 or 2 to M_t, we obtain

$$M_t \geq n_1 + 2n_2 - m_* - 2t \geq \frac{m}{2} + o(n) \geq \frac{m}{3} \ .$$

□

3.4.3 Bounding the increments

Lemma 3.4.2. *For any $v \in [n]$, any $t = O(T)$ and any $w \in [n] \setminus V_t$, we have*

$$\mathbb{P}_t(w_{t+1} = w) \leq (1 + o(1)) \frac{d_w}{M_t} \ . \tag{3.73}$$

Moreover, if $d_w = 1$

$$\mathbb{P}_t(w_{t+1} = w) \geq (1 + o(1)) \frac{d_w}{M_t} \ . \tag{3.74}$$

Proof. The proof uses an edge-switching argument. A *switching* is a local operation that transforms an input into another one. Given an input (G, Π) and two oriented edges (a, b) and (c, d) with $ab, cd \in E(G)$ and $ac, bd \notin E(G)$, we obtain the new input by deleting the edges ab and cd, and adding the edges ac and bd. Note that this operation preserves the degree of each vertex and does not modify the permutations of the adjacency lists. We will restrict to switchings that do not modify the edges within V_t in order to switch between inputs in $\mathcal{F}_t$.

Fix $w \in [n] \setminus V_t$. If $X_t = 0$, then w is chosen with probability d_w/M_t, so we may assume that $X_t > 0$. Let v_{t+1}, e_{t+1} and w_{t+1} as described in the process. Given $\mathcal{F}_t$, v_{t+1}

and e_{t+1} are fixed, while w_{t+1} is a random vertex. Let $\mathcal{A} \subseteq \mathcal{F}_t$ be the set of inputs with $w_{t+1} = w$ and $\mathcal{B} = \mathcal{F}_t \setminus \mathcal{A}$. We will estimate the number of switchings between $\mathcal{A}$ and $\mathcal{B}$ to prove the lemma.

We first proof Equation (3.73). To switch from $\mathcal{B}$ to $\mathcal{A}$, we need to switch the edges (v_{t+1}, w_{t+1}) and (w, u) for $u \in N(w)$ and there are at most d_w such switchings for each input in $\mathcal{B}$. To switch from $\mathcal{A}$ to $\mathcal{B}$, it suffices to select the edges (v_{t+1}, w) and (x, y) with $x \notin N(v_{t+1}) \cup V_t$ and $y \notin N(w)$. By Lemma 3.4.1, there are at most $\Delta\Delta' + \Delta'\Delta = o(n_1) = o(M_t)$ oriented edges (x, y) with $x \in [n] \setminus V_t$ that violate the previous condition. Thus, there are at least $(1 + o(1))M_t$ switchings for each input in $\mathcal{A}$. It follows that

$$\mathbb{P}_t(w_{t+1} = w) = \frac{|\mathcal{A}|}{|\mathcal{A}| + |\mathcal{B}|} \leq \frac{|\mathcal{A}|}{|\mathcal{B}|} \leq (1 + o(1))\frac{d_w}{M_t} .$$

We now prove Equation (3.74). Suppose that $d_w = 1$. To switch from $\mathcal{A}$ to $\mathcal{B}$ we must choose the oriented edge (v_{t+1}, w) and an oriented edge (x, y) with $x \in [n] \setminus V_t$, otherwise we would alter the edges within V_t. It follows that there are at most M_t switchings for each input in $\mathcal{A}$. To switch from $\mathcal{B}$ to $\mathcal{A}$, we must choose the oriented edge (v_{t+1}, w_{t+1}) and the unique oriented edge (w, u), where u is the only neighbour of w. Observe that if either $v_{t+1}w$ or $w_{t+1}u$ is an edge of the graph, the switching is invalid. Instead of giving a lower bound for the number of switchings of a fixed input in $\mathcal{B}$, we will give a lower bound for the average number of switchings over $\mathcal{B}$. For each $z \in [n] \setminus (V_t \cup \{w\})$, let $\mathcal{B}_z$ be the set of inputs in $\mathcal{B}$ with $w_{t+1} = z$. Given an input (G, Π) and $x \in [n] \setminus V_t$ with $d_x = 1$, we say that the input is *x-good* if $v_{t+1}x, zy \notin E(G)$, where y is the only neighbour of x; otherwise we call the input *x-bad*. Since $d_x = 1$ and $d_z \leq \Delta'$, by Lemma 3.4.1, there are at most $\Delta + \Delta'\Delta = o(n_1) = o(n_1(t))$ vertices x for which a given input is x-bad. We can generate a random input in $\mathcal{B}_z$, by first choosing one uniformly at random and then permuting the labels of the vertices of degree 1 in $[n] \setminus (V_t \cup \{z\})$. Thus, the probability that a random input in $\mathcal{B}_z$ is w-bad is $o(1)$. If an input is w-good, switching (v_{t+1}, w_{t+1})

with (w, u) yields an input in $\mathcal{A}$. It follows that

$$\mathbb{P}_t(w_{t+1} = w) = \frac{|\mathcal{A}|}{|\mathcal{A}| + |\mathcal{B}|} = \frac{1}{1 + |\mathcal{B}|/|\mathcal{A}|} \geq \frac{1}{1 + (1 + o(1))M_t} = (1 + o(1))\frac{1}{M_t}$$

□

Define $\eta_t = d_{w_t} - 2$. Next result bounds the first and second moments of η_t.

Lemma 3.4.3. *For any $v \in [n]$ and any $t = O(T)$, we have*

$$\mathbb{E}_t\left[\eta_{t+1}\right] \leq \frac{Q_0}{2} \quad \textit{and} \quad \mathbb{E}_t\left[(\eta_{t+1})^2\right] \leq 4R\ . \tag{3.75}$$

Proof. Note that $\sum_{i=1}^{t} d_{w_i}(d_{w_i} - 2) \geq -t$. Using Lemma 3.4.1 and $t = O(T) = o(|Q_0|m)$,

$$\sum_{w \in [n] \setminus V_t} d_w(d_w - 2) = \sum_{w \in [n] \setminus V_0} d_w(d_w - 2) - \sum_{i=1}^{t} d_{w_i}(d_{w_i} - 2) \leq Q_0 m + t + 1 \leq \frac{Q_0 m}{2} \leq 0\ . \tag{3.76}$$

Applying Lemma 3.4.2 and $M_t \leq m$,

$$\mathbb{E}_t\left[\eta_{t+1}\right] = \sum_{w \in [n] \setminus V_t} (d_w - 2)\mathbb{P}_t(w_{t+1} = w) \leq \frac{1 + o(1)}{M_t} \sum_{w \in [n] \setminus V_t} d_w(d_w - 2) \leq Q_0/2\ .$$

Similarly, we can bound the second moment. By Lemma 3.4.1 and Equation (3.73),

$$\mathbb{E}_t\left[(\eta_{t+1})^2\right] = \sum_{w \in [n] \setminus V_t} (d_w - 2)^2 \mathbb{P}_t(w_{t+1} = w) \leq \frac{(1 + o(1))}{M_t} \sum_{w \in [n] \setminus V_t} (d_w - 2)^2 d_w \leq 4R$$

□

3.4.4 Proof of Theorem 3.1.8

Let $\gamma := 80$. Define the stopping time

$$\tau_X = \tau_X(v) = \inf\{t : X_t = 0\} \wedge (\gamma T + 1) \ ,$$

where X_t is obtained by starting the process with $V_0 = S \cup \{v\}$. We omit the floor and ceiling functions in this section for ease of notation.

Instead of studying X_t, we focus on the stochastic process $(Z_t)_{t\geq 0}$ defined by $Z_0 = 2|Q_0|T$ and for $t \in N$

$$Z_{t+1} := Z_t + \eta_{t+1} = 2|Q_0|T + \sum_{i=0}^{t} \eta_{i+1} \ . \tag{3.77}$$

Observe that Z_t is $\mathcal{F}_t$-measurable. For any $t < \tau_X$, we can bound the increments $X_{t+1} - X_t \leq d_{w_{t+1}} - 2 = \eta_{t+1}$. Therefore, for every $t \leq \tau_X(v)$ we have $X_{t+1} \leq Z_{t+1}$.

Define the stopping time

$$\tau_Z = \tau_Z(v) := \inf\{t : Z_t = 0\} \wedge (\gamma T + 1) \ , \tag{3.78}$$

where Z_t is obtained by starting the process with $V_0 = S \cup \{v\}$. Hence, $\tau_X(v) \leq \tau_Z(v)$ and it suffices to bound the latter from above.

Write $\mu_{t+1} := (\eta_{t+1} - \mathbb{E}_t[\eta_{t+1}])\mathbb{1}t < \tau_Z$ and $S_{t+1} := \sum_{i=0}^{t} \mu_{i+1}$. For every $t < \tau_Z$, we can write

$$Z_{t+1} = 2|Q_0|T + S_{t+1} + \sum_{i=0}^{t} \mathbb{E}_i[\eta_{i+1}] \ . \tag{3.79}$$

Since $\mathbb{E}_i[\mu_{i+1}] = 0$ for all $i \geq 0$, S_t is a martingale with respect to $\mathcal{F}_t$ with $S_0 = 0$. We will use the following Bennett-type concentration inequality for martingales due to Freedman.

Lemma 3.4.4 ([29]). *Let $(S_t)_{t\geq 0}$ be a martingale with respect to a filtration $(\mathcal{F}_t)_{t\geq 0}$*

with $S_0 = 0$ *and increments* $\mu_{t+1} = S_{t+1} - S_t$. *Suppose there exists* $c > 0$ *such that* $\max_{t\geq 0} |\mu_{t+1}| \leq c$ *almost surely. For* $t \geq 0$, *define*

$$V(t+1) := \sum_{i=0}^{t} \mathbb{E}_i \left[(\mu_{i+1})^2\right] \ .$$

Then, for every $\alpha, \beta > 0$

$$\partial S_t \geq \alpha \text{ and } V(t) \leq \beta \text{ for some } t \geq 1 \leq \exp\left(\frac{-\alpha^2}{2(\beta + c\alpha)}\right) \ .$$

Deterministically, we have $\max_{t\geq 0} |\mu_{t+1}| \leq \Delta =: c$. Moreover, by Lemma 3.4.3 for all $t \geq 0$,

$$V(t) \leq \sum_{i=0}^{t-1} \mathbb{E}_i \left[(\eta_{i+1})^2 \mathbb{1} i < \tau_Z\right] \leq 4R(t \wedge \gamma T). \tag{3.80}$$

Choose $\alpha = (\gamma/3)|Q_0|T$ and $\beta = 4R\gamma T$. Thus, for all $t \geq 0$, $V(t) \leq \beta$ deterministically and, since $|Q_0| \leq 1$, $2(\beta + c\alpha) \leq 8\gamma(R + \Delta|Q_0|)T$. By Lemma 3.4.4, uniformly on the choice of $v \in [n]$

$$\mathbb{P}\left(S_t \geq \alpha \text{ for some } t\right) \leq \exp\left(-\frac{\gamma T |Q_0|^2}{72(R + \Delta|Q_0|)}\right) = O\left(1/\lambda\right) \ . \tag{3.81}$$

since $\gamma \geq 72$.

By Equation (3.75) we have $\sum_{i=0}^{\gamma T - 1} \mathbb{E}_i \left[\eta_{i+1}\right] \leq (\gamma/2)|Q_0|T$. Combining it with Equation (3.81), we obtain uniformly on $v \in [n]$

$$\partial \tau_Z(v) > \gamma T = \partial Z_t > 0 \text{ for all } t \leq \gamma T \leq \partial S_{\gamma T} > (\gamma/2 - 2)|Q_0|T = O\left(1/\lambda\right) \ . \tag{3.82}$$

since $\gamma \geq 12$.

Observe that if $|\mathcal{C}(v)| > (\gamma + 2)T$, then $\tau_Z(v) \geq \tau_X(v) > \gamma T$. As in Equation (3.67),

letting Z be the number of components of size larger than $(\gamma+2)T$ and by Equation (3.82)

$$\mathbb{E}[Z] \leq \frac{1}{(\gamma+2)T} \sum_{v\in[n]} \mathbb{P}\left(|\mathcal{C}(v)| > (\gamma+2)T\right) \leq \frac{1}{T} \sum_{v\in[n]} \mathbb{P}\left(\tau_Z(v) > \gamma T\right) = O\left(1/\log\lambda\right) = o(1)\ .$$

Markov's inequality concludes the proof.

3.5 Proof of Proposition 3.1.12

Given $\epsilon \in (0,1)$ and $\Delta = \Delta(n) = o(\sqrt{n})$ with $\log n = o(\Delta)$, define

$$\ell = \left\lfloor (1-\epsilon)\frac{n}{\Delta^2} \right\rfloor . \tag{3.83}$$

Consider the degree sequence $\widehat{\mathbf{d}}_n$ that contains $n-\ell$ vertices of degree 1 and ℓ vertices of degree Δ. We may assume that the sum of the degrees is even, otherwise we may add another vertex of degree 1. For the sake of simplicity, we will omit the floor in the definition of ℓ. Straightforward computations show that $m = (1+O(\frac{1}{\Delta}))n$ and $Q = -(1+o(1))\epsilon$.

Let $L \subseteq [n]$ denote the set of vertices of degree Δ. Let $\mathbb{G}_*$ be the random subgraph induced by $\mathbb{CM}_n(\widehat{\mathbf{d}}_n)$ on L. Let $\mathbb{G}(L,p)$ be the Erdős-Rényi random graph on the vertex set L, where each edge in $\binom{L}{2}$ is chosen independently with probability p. Let $\mathbb{P}_*(\cdot)$ and $\mathbb{P}_p(\cdot)$ be the probability measures on (multi)graphs with vertex set L associated to $\mathbb{G}_*$ and $\mathbb{G}(L,p)$, respectively, and let $\mathbb{E}_*[\cdot]$ and $\mathbb{E}_p[\cdot]$ the expected value operator defined in these probability spaces.

We briefly sketch the proof. Most of the half-edges in $\widehat{\mathbf{d}}_n$ are incident to vertices of degree 1. So typically, all vertices in L will pair most of their half-edges with the ones incident to $V \setminus L$ and the order of the largest component in $\mathbb{CM}_n(\widehat{\mathbf{d}}_n)$ will be of order at least $\Delta L_1(\mathbb{G}_*)$ whp. To estimate $L_1(\mathbb{G}_*)$, we will show that $\mathbb{G}_*$ behaves like $\mathbb{G}(L,p_*)$ with $p_* := \frac{\Delta^2}{n} = \frac{(1-\epsilon)}{\ell}$. Classic results on the subcritical regime of random graphs will give lower bounds for $L_1(\mathbb{G}(L,p_*))$ that also apply to $L_1(\mathbb{G}_*)$. We will finally use Equation (3.8) to

transfer the lower bound on the largest component from $\mathbb{CM}_n(\widehat{\mathbf{d}}_n)$ to $\mathbb{G}_n(\widehat{\mathbf{d}}_n)$.

Precisely, we will show that certain small subgraphs in $\mathbb{G}_*$ appear with the same probability as in $\mathbb{G}(L, p_*)$. Let Z_s be the number of isolated trees of size s in $\mathbb{G}_*$. Paley-Zygmund's inequality implies

$$\mathbb{P}_*(Z_s > 0) \geq \frac{\mathbb{E}_*\left[Z_s\right]^2}{\mathbb{E}_*\left[Z_s^2\right]} \ . \tag{3.84}$$

Lemma 3.5.1. *For every $s = O(\log \ell)$ we have,*

$$\mathbb{E}_*\left[Z_s\right] = (1 + o(1))\mathbb{E}_{p_*}\left[Z_s\right]$$
$$\mathbb{E}_*\left[Z_s^2\right] = (1 + o(1))\mathbb{E}_{p_*}\left[Z_s^2\right] \ .$$

Proof. Choose $S \subset L$ with $|S| = s$ and any tree T with $V(T) = S$. Let A_T be the event that S induces an isolated copy of T in L, which can be defined for $\mathbb{G}_*$ and $\mathbb{G}(L, p)$.

Fix an arbitrary ordering of $E(T)$, $e_1, \ldots, e_{s-1}$. A *realisation* of T is a set of pairs of half-edges $\{a_1b_1, \ldots, a_{s-1}b_{s-1}\}$ such that the endpoints of e_i are the vertices incident to a_i and b_i. Let $k(T)$ be the number of realisations of T. If $d_1, \ldots, d_s$ is the degree sequence of T, then $\sum_{i=1}^{s} d_i = 2(s-1)$ and

$$k(T) = \prod_{i=1}^{s} \frac{\Delta!}{(\Delta - d_i)!} = \Delta^{2(s-1)} \prod_{i=1}^{s} \left(1 + O\left(\frac{s}{\Delta}\right)\right) = (1 + o(1))\Delta^{2(s-1)} \ , \tag{3.85}$$

since $s^2 = o(\Delta)$.

The event A_T admits a partition into $k(T)$ subevents $A_T^1, \ldots, A_T^{k(T)}$ depending on the realisation of T. For $i \in [k(T)]$, $\mathbb{P}_*(A_T^i)$ is equal to the probability that $\mathbb{CM}_n$ satisfies:

(1) the i-th realisation of T is in $\mathbb{CM}_n$;

(2) for every $u \in S$ and every incident half-edge a not in the i-th realisation, a is paired in $\mathbb{CM}_n$ to a half-edge incident to $V \setminus L$.

Let $r = s(\ell - s) + \binom{s}{2}$. For $i \in [k]$, consider the i-th realisation of T, let $(a_1^i b_1^i, \ldots, a_{s-1}^i b_{s-1}^i)$

be a sequence of pairings corresponding to $E(T)$ and let $(\bar{a}_s^i\bar{b}_s^i, \ldots, \bar{a}_r^i\bar{b}_r^i)$ be a sequence of all nonpairings with at least one half-edge in S (we will assume $\bar{a}_j^i$ is always incident to S). Let B_j be the event that $a_l^i b_l^i$ is a pairing for all $l \leq j \wedge (s-1)$ and $\bar{a}_l^i\bar{b}_l^i$ is not a pairing for all $s \leq l \leq r$.

We can write

$$\mathbb{P}_*(A_T) = \sum_{i=1}^{k(T)} \prod_{j=1}^{s-1} \mathbb{P}\left(a_j^i b_j^i \in E(\mathbb{CM}_n) \mid B_{j-1}\right) \prod_{j=s}^{r} \mathbb{P}\left(\bar{a}_j^i\bar{b}_j^i \notin E(\mathbb{CM}_n) \mid B_{j-1}\right) \tag{3.86}$$

Each term on the first product in Equation (3.86) is $\frac{1}{m-O(s)} = \left(1 + O(\frac{s}{m} + \frac{1}{\Delta})\right)\frac{1}{n}$; so the first product is

$$\prod_{j=1}^{s-1} \mathbb{P}\left(a_j^i b_j^i \in E(\mathbb{CM}_n) \mid B_{j-1}\right) = \left(1 + O\left(\frac{s^2}{m} + \frac{s}{\Delta}\right)\right)\frac{1}{n^{s-1}} \tag{3.87}$$

In order to estimate the probability of (2) (which is given by the second product in Equation (3.86)), we compute the probability that each half-edge a incident to S is not paired with half-edges in L. There are exactly $s\Delta - 2(s-1)$ such events, and each has probability $1 - \frac{\ell\Delta - O(s\Delta)}{m+O(s\Delta)} = \left(1 - \frac{\ell\Delta}{n}\right)\left(1 + O(\frac{s\Delta+\ell}{n})\right)$. Thus, we have

$$\begin{aligned}
\prod_{j=s}^{r} \mathbb{P}\left(f_j f_j' \notin E(\mathbb{CM}_n) \mid B_{j-1}\right) &= \left(1 - \frac{\ell\Delta}{n}\right)^{s\Delta - 2(s-1)} \left(1 + O\left(\frac{s\Delta+\ell}{n}\right)\right)^{s\Delta} \\
&= \left(1 - \frac{\Delta^2}{n}\right)^{s\ell} e^{O(s/\ell + s/\Delta)} \left(1 + O\left(\frac{s^2}{\ell} + \frac{s}{\Delta}\right)\right) \\
&= (1 + o(1))\,(1 - p_*)^{r-s+1} \ , \qquad (3.88)
\end{aligned}$$

where we used that $(1-x/N)^y = (1-y/N)^x e^{O((x^2y+y^2x)/N^2)}$ with $x = \ell, y = \Delta, N = n/\Delta$, and that $\Delta \gg s, \ell \gg s^2$.

Plugging Equations (3.85), (3.87) and (3.88) into Equation (3.86), we obtain

$$\mathbb{P}_*(A_T) = (1 + o(1))p_*^{s-1}(1 - p_*)^{r-s+1} = (1 + o(1))\mathbb{P}_{p_*}(A_T) \ .$$

Adding over all sets $S \subset L$ with $|S| = s$ and over all trees T with $V(T) = S$, we obtain the first part of the lemma.

For the second part, choose $S, S' \subset [m]$ with $|S| = |S'| = s$ and any pair of trees T and T' with $V(T) = S$ and $V(T') = S'$. Note that $\mathbb{P}_*(A_T, A_{T'}) = 0$ unless $S = S'$ and $T = T'$, or $S \cap S' = \emptyset$. Suppose we are in the latter case, and let $r = 2s(m-2s) + \binom{2s}{2}$. Following similar computations as the ones we did for a single tree, we obtain

$$\mathbb{P}_*(A_T, A_{T'}) = (1 + o(1))\mathbb{P}_{p_*}(A_T, A_{T'}) \ ,$$

and adding over all pairs of sets and trees supported on these sets, the second part also follows. $\square$

The moments of Z_s in $\mathbb{G}(L, p)$ are well-studied in random graph theory. Let $I(\lambda) = \lambda - 1 - \ln \lambda$ be the large deviation rate function for Poisson random variables with mean $\lambda > 0$. For $\lambda = 1 - \epsilon$, any $a < (I_\lambda)^{-1}$ and $s_0 = \lfloor a \log \ell \rfloor$, we have $\mathbb{E}_{p_*}\left[Z_{s_0}^2\right] = (1 + o(1))\mathbb{E}_{p_*}\left[Z_{s_0}\right]^2$ (see e.g. Lemma 2.12(i) in [30]). Combining this with Lemma 3.5.1 and Equation (3.84), whp $\mathbb{G}_*$ has an isolated tree of size s_0. As every vertex in L has degree Δ, there are exactly $\Delta s_0 - 2(s_0 - 1)$ vertices of degree 1 that attach to the given tree. Therefore, whp there exists a component in $\mathbb{CM}_n(\widehat{\mathbf{d}}_n)$ of order $(1 + o(1))\Delta s_0$.

Observe that $I_\lambda = \frac{\epsilon^2}{2} + O(\epsilon^3)$ and since $Q = -(1 + o(1))\epsilon$, we have $I_\lambda = (1 + o(1))\frac{Q^2}{2}$. As $\mathbb{E}\left[D^2\right] = O(1)$ and $R \sim \Delta$, we can use Theorem 3.1.7 to deduce that whp

$$L_1(\mathbb{G}_n(\widehat{\mathbf{d}}_n)) \geq (1 + o(1))\Delta s_0 \geq (1 + o(1))\frac{2R}{Q^2}\log\left(\frac{n}{R^2}\right) .$$

This concludes the proof of the proposition.

Remark 3.5.2 (Concentration of $L_1(\mathbb{G}_n(\mathbf{d}_n))$)**.** Proposition 3.1.12 imposes the condition $\Delta = o(\sqrt{n})$, or equivalently $\ell \to \infty$ as $n \to \infty$. If Δ is of order $\sqrt{n}$, it is easy to check that the probability that $\mathbb{G}_* = H$ is bounded away from 0 for every H of order ℓ. Since the size of the largest component is asymptotically equal to $\Delta L_1(\mathbb{G}_*)$, $L_1(\mathbb{CM}_n(\widehat{\mathbf{d}}_n))$ and

$L_1(\mathbb{G}_n(\widehat{\mathbf{d}}_n))$ are not concentrated.

Remark 3.5.3 (The case $Q = o(1)$)**.** The largest component of Erdős-Rényi is well-studied in the barely subcritical regime (see e.g. Theorem 5.6 in [44]). If $p = \frac{1-\epsilon(\ell)}{\ell}$ with $\epsilon(\ell) > 0$ and $\ell^{-1/3} \ll \epsilon(\ell) \ll 1$, then whp

$$L_1(\mathbb{G}(\ell, p)) \sim 2\epsilon^2 \log\big(\epsilon^3 \ell\big) \ . \tag{3.89}$$

Let $\ell = (1 - \epsilon(n))\frac{n}{\Delta^2}$ and define the degree sequence $\widehat{\mathbf{d}}_n$ as before. Again, $Q \sim -\epsilon$ and $R \sim \Delta$. In particular $\Delta|Q| = o(R)$ holds.

Set $p = \frac{1-\epsilon(n)}{\ell}$. The same argument as in the proof of Proposition 3.1.12 and Equation (3.89) gives

$$L_1(\mathbb{G}_n(\widehat{\mathbf{d}}_n)) \geq \Delta L_1(\mathbb{G}(\ell, p)) \geq (1 + o(1))\frac{2\Delta}{\epsilon^2} \log\left(\frac{\epsilon^3 n}{\Delta^2}\right) \geq (1 + o(1))\frac{2R}{Q^2} \log\left(\frac{|Q|^3 n}{R^2}\right) \ . \tag{3.90}$$

The condition $\epsilon(\ell) \gg \ell^{-1/3}$ for the validity of Equation (3.89) is equivalent to $Q \leq -\omega(n) n^{-1/3} R^{2/3}$, for some $\omega(n) \to \infty$.

CHAPTER 4

BARELY SUBCRITICAL RANDOM DIGRAPHS WITH A GIVEN DEGREE SEQUENCE

4.1 Introduction

4.1.1 The directed configuration model

Let $\vec{\mathbf{d}}_n = (\mathbf{d}_n^-, \mathbf{d}_n^+) = ((d_1^-, d_1^+), \ldots, (d_n^-, d_n^+))$ be a directed degree sequence on n vertices and let $m_n = \sum d_i^- = \sum d_i^+$. The directed configuration model on $\vec{\mathbf{d}}_n$, $\mathbb{DCM} = \mathbb{DCM}(\vec{\mathbf{d}}_n)$ introduced by Cooper and Frieze [13], is the random directed multigraph on $[n]$ obtained by associating with vertex i d_1^- in-stubs and d_i^+ in-stubs, and then choosing a perfect matching of in- and out- stubs uniformly at random. This is a directed generalisation of the configuration model of Bollobás [6] which since its introduction has become one of the most widely used random graph models.

A strongly connected component in a digraph is a maximal sub-digraph such that there exists a directed path between each ordered pair of vertices. In this chapter we will consider the size of the largest strongly connected component in the barely subcritical regime.

Let $n_{k,\ell}$ be the number of copies of (k, ℓ) in $\vec{\mathbf{d}}_n$ and let $\Delta_n = \max(d_i^-, d_i^+)$ be the maximum degree of $\vec{\mathbf{d}}_n$. Also, the following are parameters of the degree distribution which govern the behaviour of the size of the largest strongly connected component.

Definition 4.1.1.

$$Q_n := \frac{1}{m_n}\sum_{i=1}^{n} d_i^- d_i^+ - 1 \quad R_n^- := \frac{1}{m_n}\sum_{i=1}^{n} d_i^- d_i^+ (d_i^- - 1) \quad R_n^+ := \frac{1}{m_n}\sum_{i=1}^{n} d_i^- d_i^+ (d_i^+ - 1)$$

We shall assume the following conditions which ensure that the degree sequence is suitably "well behaved".

Condition 4.1.2. For each n, $(d_i^-, d_i^+)_{i=1}^n = ((d_i^-, d_i^+)^{(n)})_{i=1}^n$ is a sequence of ordered pairs of non-negative integers such that $\sum_{i=1}^n d_i^- = \sum_{i=1}^n d_i^+$. Furthermore, $(p_{i,j})_{i,j=1}^\infty$ is a probability distribution such that for some $\varepsilon, \zeta > 0$,

i) $n_{i,j}/n \to p_{i,j}$ as $n \to \infty$ for each $i, j \geq 0$,

ii) $m_n/n = \mu^{(n)} \to \mu = \sum_{i,j=1}^\infty i p_{i,j} = \sum_{i,j=1}^\infty j p_{i,j}$,

iii) $n_{0,0} = 0$,

iv) $\sum_{i=1}^\infty n_{0,i} + n_{i,0} \leq (1-\varepsilon)n$,

v) $n_{1,1} \leq (1-\varepsilon)n$,

vi) $\Delta_n \leq n^{1/6}$,

vii) $R_n^-, R_n^+ \geq \zeta$.

It was shown by Cooper and Frieze [13] that $Q_n = 0$ is the threshold for the existence of a giant strongly connected component under some mild conditions on the degree sequence similar to those observed in Condition 4.1.2.

For the remainder of the chapter, we shall omit the subscript on $\vec{\mathbf{d}}_n, \Delta_n$ etc. for reasons of clarity. Let $\mathcal{C}_k(\vec{\mathbf{d}})$ be the size of the largest strongly connected component of the directed configuration model with degree sequence $\vec{\mathbf{d}}$. We shall write $\mathcal{C}_k$ for this quantity when the degree sequence is clear.

Our main result is the following,

Theorem 4.1.3. Let $\mathbb{DCM}(\vec{\mathbf{d}})$ be a configuration model random digraph with degree sequence $\vec{\mathbf{d}}$ and suppose that $nQ^3(R^-R^+)^{-1} \to -\infty$. With high probability, there are no complex components or cycles of length $\omega(1/|Q|)$. Furthermore, the probability that the kth largest cycle has length at least $\alpha|Q|^{-1}$ is

$$\mathbb{P}\left(|\mathcal{C}_k| \geq \frac{\alpha}{|Q|}\right) = 1 - \sum_{i=0}^{k-1} \frac{\xi_\alpha^i}{i!} e^{-\xi_\alpha} + o(1).$$

Where

$$\xi_\alpha = \int_\alpha^\infty \frac{e^{-x}}{x} dx.$$

In prior work, Łuczak and Seierstad [55] considered the model $D(n,p)$ which is a random digraph model formed by including each possible arc with probability p independently. They showed an analogous result to Theorem 4.1.3 for $p = (1-\varepsilon)/n$ with $\varepsilon = o(1)$ and $\varepsilon^3 n \to \infty$ in this model.

4.1.2 Previous work

The study of the giant component in random graph models was initiated by the seminal paper of Erdős and Rényi [24] regarding the giant component in $G(n,p)$. Since then the appearance of a giant component in various models has remained an active topic of study in the area.

When working with directed graphs, there are a number of types of connected component which are of interest. In this chapter we concern ourselves with the strongly connected components. The study of strongly connected components in random digraphs began in the model $D(n,p)$ where we include each possible edge with probability p independently. Karp [47] and Łuczak [53] independently showed that when $p = c/n$, then $D(n,p)$ has all strongly connected components of size $O(1)$ if $c < 1$. If instead $c > 1$ they showed that there exists a unique strongly connected component of linear order with all other components of size $O(1)$. The case $p = (1+\varepsilon)/n$ with $\varepsilon = o(1)$ and

$|\varepsilon|^3 n \to \infty$ was studied by Łuczak and Seierstad [55]. They showed that if $\varepsilon^3 n \to \infty$, there is a unique strongly connected component of size $4\varepsilon^2 n$ and other components all of size $O(1/\varepsilon)$. When $\varepsilon^3 n \to -\infty$ they showed an analogue of Theorem 4.1.3:

Theorem 4.1.4. Let $np = 1 - \varepsilon$ where $\varepsilon \to 0$ but $\varepsilon^3 n \to -\infty$. Assume $a > 0$ is a constant and let X_s denote the size of the *sth* largest strongly connected component. Then, asymptotically almost surely, $D(n,p)$ contains no complex components and

$$\lim_{n\to\infty} \mathbb{P}(X_s < a/\varepsilon) = \sum_{i=0}^{s-1} \frac{\lambda_a^i}{i!} e^{-\lambda_a},$$

where $\lambda_a = \int_a^\infty \frac{e^{-x}}{x} dx$.

The so-called critical window, when $p = (1 + \lambda n^{-1/3})/n$ has also been the subject of some study. In [15] the author showed bounds on the size of the largest strongly connected component in this regime which are akin to bounds obtained by Nachmias and Peres for G(n,p) [62]. Moreover, Goldschmidt and Stephenson [31] gave a scaling limit result for the largest strongly connected components in the critical window.

The directed configuration model has also been studied previously. It was first studied by Cooper and Frieze who showed that provided the maximum degree $\Delta < n^{1/12}/\log(n)$ then if $Q_n < 0$, there is no all strongly connected components are small and if $Q_n > 0$ there is a giant strongly connected component of linear size. The assumptions on the degree sequence have subsequently been relaxed, Graf [32] showed $\Delta < n^{1/4}$ is enough to draw the same conclusion and Cai and Perarnau [11] improved this further to only require bounded second moments. A scaling limit result was obtained at the exact point of criticality, $Q_n = 0$ by Donderwinkel and Xie [19] in a very closely related model where vertices' degrees are sampled from a limiting degree sequence.

4.1.3 Organization

The remainder of this chapter is arranged as follows, in Section 4.2 we prove some auxiliary results on subgraph counting within the directed configuration model. Then, in Section 4.3 we enumerate certain types of strongly connected directed graphs of maximum degree 4. In section Section 4.4 we prove Theorem 4.1.3 which we break down into a few steps, first that there are no long cycles. Next that there are no complex components and finally we compute the probability the kth largest component has size at least $\alpha/|Q|$ via Poisson approximation. We conclude the chapter in Section 4.5 with some open questions and future work.

4.2 Subgraph Bounds

We calculate probabilities of given subgraphs in the configuration model. We let (D_-, D_+) be the degree distribution, that is the random variable obtained by picking an element of $\vec{\mathbf{d}}$ uniformly at random.

Suppose that $\mathbb{DCM}$ has degree distribution (D_-, D_+) define

$$\mu_{i,j} = \mathbb{E}(D_-^i D_+^j), \qquad \text{and} \qquad \rho_{i,j} = \mathbb{E}\Big(\prod_{k=0}^{i-1}(D_- - k)\prod_{\ell=0}^{j-1}(D_+ - \ell)\Big).$$

So that the $\mu_{i,j}$ and $\rho_{i,j}$ are the moments and factorial moments of (D_-, D_+) respectively. We also define $\mu = \mu_{1,0} = \mu_{0,1}$ which is the average degree. Furthermore, observe that $\mu_{1,1} = \frac{m}{n}(1+Q)$ which is a fact we utilise in subsequent sections. We now state a general upper bound on the probability of finding certain subgraphs in the configuration model.

Lemma 4.2.1. *Let $\mathbb{DCM}$ be a configuration model random digraph with degree distribution (D_-, D_+). Suppose further that $\mathbb{DCM}$ has n vertices and m edges. Let H be any digraph with h vertices, k edges and degree sequence $H = (h_{i,j} : i, j \in \mathbb{N})$. Then the probability that a uniformly random injective map $\phi : V(H) \to V(\mathbb{DCM})$ is a homomorphism*

is bounded above by

$$p^+(H, \mathbb{DCM}) := \frac{n^h}{(n)_h (m)_k} \prod_{i,j \in \mathbb{N}} \rho_{i,j}^{h_{i,j}} \tag{4.1}$$

Furthermore, the expected number of copies of H in $\mathbb{DCM}$ is bounded above by

$$n^+(H, \mathbb{DCM}) := \frac{h!}{aut(H)} \binom{n}{h} p^+(H, \mathbb{DCM}). \tag{4.2}$$

Proof. Note that (4.2) follows immediately from (4.1). Thus we shall focus on the proof of (4.1). Let $\psi : V(H) \to V(\mathbb{DCM})$ be a fixed injective map. Arbitrarily order the edges of H as $E_1, E_2, \ldots, E_k$ and for each edge E_i define an event $\mathcal{E}_i := \{\psi(E_i) \in E(\mathbb{DCM})\}$. Then, ψ is a homomorphism if and only if every event $\mathcal{E}_1, \ldots, \mathcal{E}_k$ occurs. To simplify notation, let $\mathcal{F}_1 = \mathcal{E}_1$ and $\mathcal{F}_i = \mathcal{E}_i | \cap_{j=1}^{i-1} \mathcal{E}_j$ for $i \geq 2$ and note that

$$\bigcap_{i=1}^{k} \mathcal{E}_i = \bigcap_{i=1}^{k} \mathcal{F}_i.$$

Suppose that $E_i = a_i b_i$ for some $a_i, b_i \in V(H)$. Define $s_i = |\{j < i : a_i = a_j\}|$ and $t_i = |\{j < i : b_i = b_j\}|$. That is, s_i and t_i are the number of times that a_i (resp. b_i) has previously appeared as the initial (terminal) vertex of an edge of H. Also, suppose that $\psi(E_i) = a_i' b_i'$.

Claim.

$$\mathbb{P}(\mathcal{F}_i) \leq \frac{\min(d^+(a_i') - s_i, 0) \min(d^-(b_i') - t_i, 0)}{m + 1 - i}$$

Proof. To see this, note that if $d^+(a_i') \leq s_i$, then $\mathbb{P}(\mathcal{F}_i) = 0$ as there are not enough stubs at a_i' to create such a copy of H. Similarly if $d^+(b_i') \leq t_i$, $\mathbb{P}(\mathcal{F}_i) = 0$.

So, without loss of generality we may assume that $(d^+(a_i') - s_i), (d^-(b_i') - t_i) > 0$. Now, by definition of $\mathcal{F}_i$ we have already chosen $i - 1$ edges of the configuration model and we have precisely $d^+(a_i') - s_i$ out-stubs remaining at a_i' and $d^-(b_i') - t_i$ at b_i'. Now, consider the random matching on the $2(m + 1 - i)$ remaining stubs. The probability that we contain any given $a_i' - b_i'$ edge is $(m + 1 - i)^{-1}$. There are $(d^+(a_i') - s_i)(d^-(b_i') - t_i)$ such potential

edges. Thus the expected number of $a_i' - b_i'$ edges is $(d^+(a_i') - s_i)(d^-(b_i') - t_i)/(m+1-i)$. The claim then follows by Markov's inequality. $\square$

Next we compute the probability that all of the $\mathcal{F}_i$ occur simultaneously. Due to the way in which we defined these events,

$$\mathbb{P}(\psi \text{ is a homomorphism}) = \mathbb{P}\left(\bigcap_{i=1}^{k} \mathcal{F}_i\right) = \prod_{i=1}^{k} \mathbb{P}(\mathcal{F}_i).$$

To write down this probability succinctly, we will use the functions

$$f_{a,b}(x,y) := \prod_{i=1}^{a}\prod_{j=1}^{b}(x+1-i)(y+1-j)$$

It is a simple computation to check that

$$\mathbb{P}\left(\bigcap_{i=1}^{k} \mathcal{F}_i\right) \leq ((m)_k)^{-1} \prod_{v \in V(H)} f_{d_H^-(v), d_H^+(v)}(d^-_{\mathbb{DCM}}(\psi(v)), d^+_{\mathbb{DCM}}(\psi(v))) \tag{4.3}$$

Note that here we were able to remove the min function as if there are any negative contributions to the product, then there is also a contribution of value 0 and furthermore, there it is impossible for ψ to be a homomorphism in this case so that (4.3) reduces to $0 \leq 0$ in this case (which is clearly true).

To complete the proof, we extend the right hand side of equation (4.3) to allow any function $\psi : V(H) \to V(\mathbb{DCM})$. Choosing an uniformly random function in this way gives that the probability that a uniformly random injective function is a homomorphism is bounded above by

$$\frac{1}{(n)_h (m)_k} \sum_{\psi : V(H) \hookrightarrow V(G)} \prod_{v \in V(H)} f_{d_H^-(v), d_H^+(v)}(d^-_{\mathbb{DCM}}(\psi(v)), d^+_{\mathbb{DCM}}(\psi(v))) \tag{4.4}$$

$$\leq \frac{1}{(n)_h (m)_k} \sum_{\psi : V(H) \to V(G)} \prod_{v \in V(H)} f_{d_H^-(v), d_H^+(v)}(d^-_{\mathbb{DCM}}(\psi(v)), d^+_{\mathbb{DCM}}(\psi(v))) \tag{4.5}$$

In moving from injective functions to arbitrary functions between (4.4) and (4.5), we

move from an intractable space of functions to a product space which naturally splits over the vertices of H. This allows us to rewrite (4.5) as

$$\begin{aligned}
&\frac{1}{(n)_h(m)_k}\sum_{\psi:V(H)\to V(G)}\prod_{v\in V(H)} f_{d_H^-(v),d_H^+(v)}(d^-_{\mathbb{DCM}}(\psi(v)),d^+_{\mathbb{DCM}}(\psi(v)))\\
=&\frac{1}{(n)_h(m)_k}\prod_{v\in V(H)}\sum_{w\in V(G)} f_{d_H^-(v),d_H^+(v)}(d^-_{\mathbb{DCM}}(w),d^+_{\mathbb{DCM}}(w))\\
=&\frac{n^h}{(n)_h(m)_k}\prod_{v\in V(H)}\rho_{d_H^-(v),d_H^+(v)}=\frac{n^h}{(n)_h(m)_k}\prod_{i,j\in\mathbb{N}}\rho_{i,j}^{h_{i,j}}.
\end{aligned}$$

Which is the claimed upper bound, $p^+(H,\mathbb{DCM})$. □

We also need a lower bound on the probability that $\mathbb{DCM}$ contains a cycle.

Lemma 4.2.2. *Let $\mathbb{DCM}$ be the configuration model random digraph with degree distribution (D_-,D_+) and maximum degree Δ. Suppose further that $\mathbb{DCM}$ has n vertices and m edges. Let H be a directed cycle with h vertices. Then the probability that a uniformly random injective map $\phi:V(H)\to V(\mathbb{DCM})$ is a homomorphism is bounded from below by*

$$p_c^-(H,\mathbb{DCM}):=\frac{(1+Q)^h}{(n)_h}\left(1-\frac{2h^2\Delta^2}{\varepsilon n}\right)\left(1-\frac{h\Delta^2}{2m}\right). \tag{4.6}$$

Proof. First, let us consider the probability of finding at least one edge from vertex u of out-degree a to vertex v with in-degree b. We assume that we have not observed any edges that affect either the out-degree of u or the in-degree of v and that $\mathbb{DCM}$ has m edges. So let X be the number of edges between u and v and $\mathcal{H}_k$ our knowledge of $\mathbb{DCM}$ up until now which consists only of a set of k edges which are present. Then,

$$\begin{aligned}
\mathbb{P}(X=0|\mathcal{H}_k)&=\left(1-\frac{a}{m-k}\right)\left(1-\frac{a}{m-k-1}\right)\dots\left(1-\frac{a}{m+1-k-b}\right)\\
&\le\left(1-\frac{a}{m}\right)^b\\
&\le 1-\frac{ab}{m}+\frac{a^2b^2}{2m^2}.
\end{aligned}$$

Where the final line follows by the inequality, $(1+x)^n\le 1-nx+\binom{n}{2}x^2$ which is valid

for $n \in \mathbb{N}$ and $x \geq -1$. This allows us to deduce that the probability of seeing an edge between u and v is at least $\frac{ab}{m} - \frac{a^2b^2}{2m^2}$. Now, looking at each edge of H in turn noting allows us to deduce that the probability ϕ is a homomorphism is at least

$$\frac{1}{(n)_h} \sum_{\phi:[h]\hookrightarrow[n]} \prod_{i=1}^{h} \left(\frac{d^-_{\phi(i)} d^+_{\phi(i+1)}}{m} - \frac{(d^-_{\phi(i)} d^+_{\phi(i+1)})^2}{2m^2} \right) \tag{4.7}$$

Where we consider the argument of ϕ modulo h in (4.7). Also, note that one can factorise the linear term in 4.7 and bound the second occurrence of $d^-_{\phi(i)} d^+_{\phi(i+1)}$ by Δ^2 to get the lower bound

$$\frac{1}{(n)_h} \sum_{\phi:[h]\hookrightarrow[n]} \prod_{i=1}^{h} \frac{d^-_{\phi(i)} d^+_{\phi(i)}}{m} \left(1 - \frac{\Delta^2}{2m} \right) \tag{4.8}$$

The idea is to argue that we can swap the order of the product and sum in Equation (4.8) without changing the result very much. To this end, let Φ_i be the set of functions $\phi : [h] \to [n]$ such that

i) $|\phi([n])| = h - i$,

ii) $d^+_{\phi(j)} \neq 0$ for each $j \in [h]$,

iii) $d^-_{\phi(j)} \neq 0$ for each $j \in [h]$.

Note that for any function $\phi \notin \bigcup_{i=0}^{h} \Phi_i$ then $\prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)} = 0$. As a result of this we observe that

$$\sum_{\phi:[h]\hookrightarrow[n]} \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)} = \sum_{\phi\in\Phi_0} \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)}.$$

As well as the fact that

$$\sum_{j=0}^{h} \sum_{\phi\in\Phi_j} \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)} = \sum_{\phi:[h]\to[n]} \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)} = \left(\sum_{i=1}^{n} d^-_i d^+_i \right)^h = m^h (1+Q)^h.$$

For a given function $\phi : [h] \to [n]$ we define its weight as $w(\phi) := \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)}$. We also define the weight of a set S of functions in the natural way as $w(S) = \sum_{\phi\in S} w(\phi)$. Next,

we shall apply switching arguments to bound the weights of the sets Φ_i relative to one another.

Consider the auxiliary bipartite graphs, $\mathcal{G}_i$ with parts Φ_i and Φ_{i+1}. We connect $\phi_i \in \Phi_i$ to $\phi_{i+1} \in \Phi_{i+1}$ by an edge in $\mathcal{G}_i$ if ϕ_i and ϕ_{i+1} differ in precisely one coordinate. For each $\phi_i \in \Phi_i$ there are at most h^2 ways we can change one coordinate and decrease the size of the image. In particular we may pick any of the h coordinates of ϕ_i and change it to $\phi_i(j)$ for some $j \in [h]$. So, $\Delta_{\mathcal{G}_i}(\Phi_i) \leq h^2$. For each $\phi_{i+1} \in \Phi_{i+1}$, we pick any of the at least i coordinates at which ϕ_{i+1} is not injective and choose a new image for this coordinate. By Condition 4.1.2 there are at least $\varepsilon n/2$ ways to choose the new image. $\delta_{\mathcal{G}_i}(\Phi_{i+1}) \geq i\varepsilon n/2$.

Combining these two results allows us to deduce that $i\varepsilon n/2|\Phi_{i+1}| \leq e(\mathcal{G}_i) \leq h^2|\Phi_i|$. Upon rearrangement we find $|\Phi_i| \geq \frac{\varepsilon n}{2h^2}|\Phi_{i+1}|$. Note that two functions ϕ and ψ which differ in one coordinate must also satisfy $w(\phi) \leq \Delta^2 w(\psi)$ and vice versa. Hence $w(\Phi_i) \geq \frac{\varepsilon n}{2h^2\Delta^2} w(\Phi_{i+1})$. This allows us to apply induction to deduce that

$$\sum_{j=0}^{h} \sum_{\phi \in \Phi_j} \prod_{i=1}^{h} d^-_{\phi(i)} d^+_{\phi(i)} = \sum_{j=0}^{h} w(\Phi_j) \leq w(\Phi_0) \sum_{j=0}^{h} \left(\frac{2h^2\Delta^2}{\varepsilon n} \right)^j \leq \frac{w(\Phi_0)}{1 - \frac{2h^2\Delta^2}{\varepsilon n}}.$$

So, $w(\Phi_0) \geq (1 - \frac{2h^2\Delta^2}{\varepsilon n}) m^h (1+Q)^h$. Combining this with (4.8) allows us to deduce the statement of the lemma, that the following is a lower bound on the probability of finding a cycle at a specified position in $\mathbb{DCM}$:

$$\frac{(1+Q)^h}{(n)_h} \left(1 - \frac{2h^2\Delta^2}{\varepsilon n} \right) \left(1 - \frac{h\Delta^2}{2m} \right).$$

□

4.3 Digraph Counting

In this section we will prove the following bound on the number of strongly connected multi-digraphs with maximum degree 4. This is similar to [15, Lemma 2.3] although we a weaker bound suffices here allowing us to simplify the proof somewhat.

Lemma 4.3.1. *Suppose that $n, m, a, b \in \mathbb{N}$ such that $n + a + b = m$. Let $N(n, a, b)$ be the number of labelled strongly connected multi-digraphs with n vertices and degree distribution given by*

in-degree	*out-degree*	*quantity*
1	1	$n - 2a - b$
1	2	a
2	1	a
2	2	b

Then, we have the following bound,

$$N(n, a, b) \leq (3a + 2b)(m - 1)! \binom{n}{a, a, b}$$

To prove this bound we will use the preheart configuration model of Pérez-Giménez and Wormald [66] which we shall define as follows.

A *preheart* is a multi-digraph with minimum semi-degree at least 1 and no cycle components. The *heart* of a preheart D is the multidigraph $H(D)$ formed by suppressing all vertices of D which have in and out degree precisely 1. For a degree sequence $\vec{\mathbf{d}}$, define

$$T = T(\vec{\mathbf{d}}) = \{v \in V : d^+(v) + d^-(v) \geq 3\}.$$

To form the *preheart configuration model*, first we apply the configuration model to T to produce a heart H. Given a heart configuration H, we construct a preheart configuration Q by assigning $V \setminus T$ to $E(H)$ such that the vertices assigned to each arc of H are given a linear order. Denote this assignment including the orderings by q. Then the preheart

configuration model, $\mathcal{Q}(\vec{\mathbf{d}})$ is the probability space of random preheart configurations formed by choosing H and q uniformly at random.

It is easy to see that every strongly connected digraph is produced by the preheart configuration model. Thus counting the number of possible outcomes from the preheart configuration model gives an upper bound for the number of strongly connected digraphs with the same degree sequence. We now count the number of preheart configurations using [15, Lemma 2.4].

Lemma 4.3.2. *In the preheart configuration model with n vertices, m edges and degree sequence $\vec{\mathbf{d}}$. Let $n' = |T(\vec{\mathbf{d}})|$ be the number of vertices of the heart. Then there are a total of*

$$\frac{n' + m - n}{m} m!$$

preheart configurations.

From this lemma, we may prove Lemma 4.3.1.

Proof of Lemma 4.3.1. First, we choose the degree sequence. So note that there are at most $\binom{n}{a,a,b}$ ways in which we can give a vertices in-degree 1 and out-degree 2, a vertices in-degree 2 and out-degree 1, b vertices in-degree 2 and out-degree 2 and the remainder in-degree 1 and out-degree 1. Having fixed this degree sequence, by Lemma 4.3.2 as the heart contains $2a + b$ vertices and there are $a + b$ more edges than vertices, the number of strongly connected digraphs with this degree sequence is bounded above by $\frac{3a+2b}{m} m!$. Hence the number of strongly connected digraphs with degree distribution as in the statement of the lemma is at most

$$(3a + 2b)(m - 1)! \binom{n}{a, a, b}$$

as claimed. □

4.4 Proof of Theorem 4.1.3

In this section we prove Theorem 4.1.3 and show that every strongly connected component is a cycle and that these cycles are not particularly large. In particular this provides a configuration model analogue of [55, Theorem 7]. Working with the configuration model introduces several additional difficulties with the proof of this, foremost of these being that we do not have enough control of the subgraph counts to compute the factorial moments and show that they converge to those of a Poisson distribution. We will instead use the Chen-Stein method for Poisson approximation which only requires good control of the first moment and an upper bound on the second.

The proof of this theorem splits naturally into four parts. For functions $f(n) \gg g(n)$ which are defined such that $f(n) = \omega(\sqrt{m/|Q|})$, $f(n) = o(m|Q|/R^-)$ and $g(n) = \omega(1/|Q|)$. Moreover for this section we shall assume that $R^- \geq R^+$ and if this is not the case, we swap the orientations of all edges to get an equivalent digraph with $R^- \geq R^+$ as desired. We will say that a cycle C is

- *Long* if $|C| \geq f(n)$,
- *Medium* if $g(n) < |C| < f(n)$,
- *Short* if $|C| \leq g(n)$.

First we will show that there are no long or medium cycles in the directed configuration model. Next, we show that there are no complex components and finally we show the result on the distribution of the length of the kth longest cycle.

4.4.1 Long Cycles

Lemma 4.4.1. $\mathbb{DCM}$ *has no long cycles.*

To show that there are no long cycles, it suffices to show that the out-component of an arbitrary vertex is bounded above by $f(n)$. Certainly, the longest cycle in a directed graph is at most the size of the largest out-component and so the lemma follows.

Thus, consider the following version of the branching process of Hatami and Molloy [35] for the out-component in a digraph. For a vertex v we explore its out-component in $\mathbb{DCM}$ as follows. We will have a partial subdigraph C_t at time t consisting of the vertices explored thus far. C_t will consist of all in- and out-stubs of some vertices of $\mathbb{DCM}$ together with a matching of some of the stubs. If there are unmatched out-stubs in C_t we will pick one at random and match it to some in-stub which yields an edge of C_t. We define Y_t as the number of unmatched out-stubs in C_t. Thus, $Y_t = 0$ indicates that we have explored an out-component in its entirety. Formally we define the exploration process as follows.

- Choose a vertex v and initialise $C_0 = \{v\}$ and $Y_0 = d^+(v)$.
- While $Y_t > 0$, choose an arbitrary unmatched out-stub of any vertex $v \in C_t$. Pick a uniformly random unmatched in-stub and let u be the vertex to which this in-stub belongs. Match these two stubs forming an edge of $\mathbb{DCM}$.
  - If $u \notin C_t$ we add it so $C_{t+1} = C_t \cup \{u\}$ and $Y_{t+1} = Y_t + d^+(u) - 1$.
  - Otherwise, $C_{t+1} = C_t$, $Y_{t+1} = Y_t - 1$.

Note that this does not depend on how we have exposed C_t so C_t and Y_t are Markov processes. We define the following quantities.

- $D_t := Y_t + \sum_{u \notin C_t} d^+(u)$, the number of unmatched out-stubs at time t. Note this is also the number of unmatched in-stubs at time t.
- $v_t := \emptyset$ if C_{t-1} and C_t have the same vertex set. Otherwise it is the unique vertex in $C_t \setminus C_{t+1}$.
- $Q_t := \frac{\sum_{u \notin C_t} d^-(u) d^+(u)}{D_t} - 1$.

Note that initially $Q_t = Q$. Also, for unvisited vertices $u \notin C_t$, the probability that we explore u next is $\mathbb{P}(v_{t+1} = u) = \frac{d^-(u)}{D_t}$. Hence, provided that $Y_t > 0$, the expected change

in Y_t is

$$\mathbb{E}(Y_{t+1} - Y_t | C_t) = \sum_{u \notin C_t} \mathbb{P}(v_{t+1} = u)(d^+(u) - 1) = \frac{\sum_{u \notin C_t} d^-(u) d^+(u)}{D_t} - 1 = Q_t. \quad (4.9)$$

As long as Q_t remains close to Q we expect that Y_t is a random walk with drift approximately Q. So in particular, for our setting of $Q < 0$, we expect that the random walk will quickly return to 0. Thus, we shall start by showing that the drift parameter Q_t is indeed close to Q with high probability. For this we shall use the following formulation of the Azuma-Hoeffding inequality (see [44, Theorem 2.25]).

Theorem 4.4.2 (Azuma-Hoeffding Inequality)**.** Let $(X_k)_{k=0}^n$ be a martingale with $X = X_n$, $X_0 = \mathbb{E}(X)$. Suppose there exist constants, $c_k > 0$ such that

$$|X_k - X_{k-1}| \leq c_k \text{ for all } k \leq n.$$

Then for any $\lambda \geq 0$,

$$\mathbb{P}(|X_n - X_0| \geq \lambda) \leq 2 \exp\left(-\frac{\lambda^2}{2\sum_{k=1}^n c_k^2}\right).$$

For $t \geq 1$ define $W_t := Q_t - Q_{t-1} - \mathbb{E}(Q_t - Q_{t-1} | C_{t-1})$. Also, we define $X_0 = Q$ and for $t \geq 1$ let

$$X_t := X_0 + \sum_{i=1}^t W_i = Q_t - \sum_{i=1}^t \mathbb{E}(Q_i - Q_{i-1} | C_{i-1}). \quad (4.10)$$

It is a simple check that the X_t form a martingale. Furthermore, $|Q_t - Q| \leq |Q_t - X_t| + |X_t - Q|$ so to bound the probability that $|Q_t - Q|$ is large, we show that $|Q_t - X_t|$ is small and bound the probability that $|X_t - Q|$ is large.

For the second of these, consider the auxiliary random variables

$$\widetilde{Q}_t := \frac{\sum_{u \notin C_t} d^-(u) d^+(u)}{D_{t-1}} - 1.$$

That is we change Q_t to have the same denominator as Q_{t-1}. As we assume that $t \leq m/2$,

$D_t \geq m/2$ so combining this with the fact that $|D_t - D_{t-1}| \leq 1$ we deduce that

$$|Q_t - \widetilde{Q}_t| = \frac{|D_t - D_{t-1}| \sum_{u \notin C_t} d^-(u) d^+(u)}{D_t D_{t-1}} \leq \frac{4}{m}. \tag{4.11}$$

Also, as there is at most one vertex whose contributions are removed in moving from Q_{t-1} to Q_t, then

$$|Q_{t-1} - \widetilde{Q}_t| \leq \frac{4\Delta^2}{m}. \tag{4.12}$$

Combining equations (4.11) and (4.12) gives an upper bound on $|Q_t - Q_{t-1}|$ which we may then use to bound the martingale differences almost surely,

$$|X_t - X_{t-1}| \leq \frac{8\Delta^2 + 8}{m} \leq \frac{16\Delta^2}{m}. \tag{4.13}$$

Next, we will bound the terms $\mathbb{E}(Q_t - Q_{t-1}|C_{t-1})$. It will be convenient to do this in two stages utilising the auxiliary random variables $\widetilde{Q}_t$. So, first note that

$$\begin{aligned} 0 \leq \mathbb{E}(Q_{t-1} - \widetilde{Q}_t | C_{t-1}) &= \sum_{u \notin C_{t-1}} \mathbb{P}(v_t = u) \frac{d^-(u) d^+(u)}{D_{t-1}} \leq \sum_{u \in V(G)} \frac{(d^-(u))^2 d^+(u)}{D_{t-1}^2} \\ &\leq \sum_{u \in V(G)} \frac{4(d^-(u))^2 d^+(u)}{m^2} \leq \frac{4R^- + 4}{m}. \end{aligned} \tag{4.14}$$

We can combine (4.14) with (4.11) to deduce that

$$|\mathbb{E}(Q_t - Q_{t-1}|C_{t-1})| \leq \frac{4R^- + 8}{m} \leq \frac{12R^-}{\zeta m}. \tag{4.15}$$

This leaves us in a situation in which we can compare X_t and Q_t,

$$|X_t - Q_t| \leq \sum_{i=1}^{t} |\mathbb{E}(Q_i - Q_{i-1}|C_{i-1})| \leq \frac{12R^- t}{\zeta m}. \tag{4.16}$$

So provided that $t \leq \frac{m\zeta|Q|}{48R^-}$ we have $|X_t - Q_t| \leq |Q|/4$ and in particular, for any such t,

$$\mathbb{P}\left(Q_t - Q \geq \frac{|Q|}{2}\right) = \mathbb{P}\left(Q_t - X_t + X_t - Q \geq \frac{|Q|}{2}\right)$$
$$\leq \mathbb{P}\left(X_t - Q \geq \frac{|Q|}{4}\right) \leq \mathbb{P}\left(|X_t - Q| \geq \frac{|Q|}{4}\right)$$

Whereupon we can apply the Azuma-Hoeffding inequality as $X_0 = Q$. We can use the bound from (4.13) for the c_k. Substituting into Theorem 4.4.2 yields,

$$\mathbb{P}\left(|X_t - Q| \geq \frac{|Q|}{4}\right) \leq \exp\left(-\frac{|Q|^2 m^2}{8192 t \Delta^4}\right) \leq \exp\left(-C|Q|mn^{-2/3}\log(n)\right). \tag{4.17}$$

Where the second inequality in (4.17) comes from $t \leq \frac{m\zeta|Q|}{48R^-}$, $\Delta \leq n^{1/6}\log^{-1/4}(n)$ and $R^- \geq \zeta$. We could improve the dependence on Δ by using Freedman's inequality [29] in place of the Azuma-Hoeffding inequality here however there are other points where we require $\Delta \leq n^{1/6}$ and so this would only remove the $\log^{-1/4}(n)$ term in Condition 4.1.2. Note $mn^{-2/3} \geq m^{1/3}/2$ hence $|Q|mn^{-2/3} \to \infty$ and so for any large enough n,

$$\mathbb{P}\left(Q_t - Q \geq \frac{|Q|}{2}\right) = \mathbb{P}\left(Q_t \geq -\frac{|Q|}{2}\right) \leq n^{-2}. \tag{4.18}$$

Now that we have shown that Q_t is concentrated around Q, we can proceed to show that $\mathbb{DCM}$ has no large components with high probability via a stopping time argument. We will use the following version of Doob's optional stopping theorem [77, Theorem 10.10]

Theorem 4.4.3 (Optional Stopping Theorem)**.** Let X be a supermartingale and let τ be a stopping time. Then X_τ is integrable and furthermore,

$$\mathbb{E}(X_\tau) \leq \mathbb{E}(X_0)$$

whenever τ is bounded.

So, now let us show that there is no component of size larger than $f(n)$. For each $v \in V(G)$ we shall consider the exploration process started at v. Recall that $f(n) =$

$o(n|Q|/R^-)$ and so for sufficiently large n, $f(n) \leq \frac{m\zeta|Q|}{48R^-}$. Hence we have $Q_t \leq -|Q|/2$ with high probability for each $t \leq f(n)$. Define the stopping time

$$\tau := \min\{t \geq 0 | Y_t = 0 \text{ or } Q_t \geq -|Q|/2 \text{ or } t = f(n)\}$$

Recall that for $t \leq \frac{m\zeta|Q|}{48R^-}$ we have $\mathbb{E}(Y_t - Y_{t-1}) = Q_{t-1} \leq -|Q|/2$. Thus, $Y_{\min(t,\tau)} + |Q|\min(t,\tau)/2$ is a supermartingale. Clearly, τ is bounded by $f(n)$ so we may apply Theorem 4.4.3 to $Y_{\min(t,\tau)} + |Q|\min(t,\tau)/2$ from which we deduce that

$$\mathbb{E}\left(Y_\tau + \frac{|Q|\tau}{2}\right) \leq Y_0 = d^+(v).$$

Upon rearrangement this yields,

$$\mathbb{E}(\tau) \leq 2\frac{d^+(v) - \mathbb{E}(Y_\tau)}{|Q|} \leq \frac{2d^+(v)}{|Q|}.$$

By Markov's inequality we can deduce

$$\mathbb{P}(\tau = f(n)) \leq \frac{2d^+(v)}{|Q|f(n)}$$

The only other way in which we could have $Y_\tau \neq 0$ is if for some i we have $Q_i \geq -|Q|/2$. A union bound allows us to deduce that this occurs with probability at most $f(n)n^{-2} \leq n^{-1}$. So for any large enough n,

$$\mathbb{P}(Y_\tau \neq 0) \leq \frac{2d^+(v)}{|Q|f(n)} + \frac{1}{n}.$$

Define Z as the number of vertices of $\mathbb{DCM}$ which lie in cycles of size at least $f(n)$. Note

that any such vertex must have out component of size at least $f(n)$. Thus,

$$\mathbb{P}(|\mathcal{C}_1| \geq f(n)) = \mathbb{P}(Z \geq f(n)) \leq \frac{\mathbb{E}(Z)}{f(n)} \leq \frac{1}{f(n)} \sum_{v \in V(\mathbb{DCM})} \mathbb{P}(|C^+(v)| \geq f(n))$$
$$\leq \frac{1}{f(n)} \sum_{v \in V(\mathbb{DCM})} \left(\frac{2d^+(v)}{|Q|f(n)} + \frac{1}{n} \right) = \frac{m}{|Q|f(n)^2} + \frac{1}{f(n)} = o(1)$$

Thus there are no long cycles.

4.4.2 Medium Cycles

Lemma 4.4.4. $\mathbb{DCM}$ *has no medium cycles.*

Our next step is to apply Lemma 4.2.1 to show there are no medium cycles. By Lemma 4.2.1, the probability that $\mathbb{DCM}$ has a cycle of length h in any particular location is at most $\frac{n^h \mu_{1,1}^h}{(n)_h (m)_h}$. Thus the expected number of cycles of length h in G is at most

$$\frac{h!}{|\mathrm{aut}(\overrightarrow{C_h})|} \binom{n}{h} \frac{n^h \mu_{1,1}^h}{(n)_h (m)_h} = \frac{m^h}{(m)_h} \frac{(1+Q)^h}{h} \leq \left(1 + \frac{h}{m-h}\right)^h \frac{(1+Q)^h}{h}$$
$$\leq \left(1 + \frac{2h}{m}\right)^h \frac{(1+Q)^h}{h} \leq \frac{e^{hQ + \frac{2h^2}{m}}}{h}. \tag{4.19}$$

For any $g(n) \leq h \leq f(n)$, we have $2h^2/m \leq 4h^2/n \leq h|Q|/2$ as $f(n) = o(n|Q|)$. Then, the expected number of cycles of length between $g(n)$ and $f(n)$ is at most

$$\sum_{h=g(n)}^{f(n)} \frac{e^{hQ + \frac{2h^2}{m}}}{h} \leq \sum_{h=g(n)}^{f(n)} \frac{e^{\frac{hQ}{2}}}{h} \leq \int_{g(n)}^{\infty} \frac{e^{\frac{hQ}{2}}}{h} dh = \int_{-\frac{Qg(n)}{2}}^{\infty} \frac{e^{-\lambda}}{\lambda} d\lambda = E_1\left(-\frac{Qg(n)}{2} \right). \tag{4.20}$$

Where $E_1(x)$ is the exponential integral function and the first equality follows by making the substitution $\lambda = -Qh/2$ (recall that $Q < 0$ and so this substitution preserves positivity). It is straightforward to bound $0 \leq E_1(x) \leq e^{-x}/x$ which allows us to conclude that $E_1(x) \to 0$ as $x \to \infty$. Note that $g(n) = \omega(1/|Q|)$ and so $-Qf(n) \to \infty$. Thus the expected number of cycles in $\mathbb{DCM}$ of length at least $g(n)$ is $o(1)$. So by Markov's inequality, there are no such cycles with high probability.

4.4.3 Complex Components

Lemma 4.4.5. $\mathbb{DCM}$ *has no complex components.*

We begin by defining digraphs $S(a,b,c)$ and $T(a,b)$ for $a,b,c \in \mathbb{N}$. Let $S(a,b,c)$ be the digraph with $a+b+c-1$ vertices consisting of vertices u, v and three internally disjoint paths. One of length a from u to v, one of length b from u to v and one of length c from v to u. Let $T(a,b)$ be the digraph with $a+b-1$ vertices consisting of two cycles, one of length a, one of length b which intersect at a single vertex, u. We can use the ear decomposition of a strongly connected digraph to deduce that if $\mathbb{DCM}$ contains any complex components, then it contains a subgraph which is either a copy of $S(a,b,c)$ or a copy of $T(a,b)$. Note that both of these are the union of two cycles and and by the results of the previous two sections, there are no cycles with more than $g(n)$ vertices with high probability. Thus we only need to show there are none of these motifs on at most $2g(n)$ vertices to deduce that there are none in $\mathbb{DCM}$.

Unlike Łuczak and Seierstad [55], we must treat these cases separately as in the first case, we have two vertices of degree 3 and the rest of degree 2 and in the second there is one vertex of degree 4 in place of the degree 3 vertices which changes the result of applying Lemma 4.2.1.

First let us consider $S(a,b,c)$. There are at most $h^2h!$ ways of finding such subgraphs on h vertices ($\leq h^2$ ways of choosing path lengths connecting the two degree 3 vertices and assuming the associated automorphism groups are all trivial gives this bound). Thus, we may apply Lemma 4.2.1 to deduce that the expected number of such subgraphs in the configuration model with parameters as in the statement of Theorem 4.1.3 is at most

$$\sum_{h=1}^{2g(n)} \frac{n^h}{(m)_{h+1}} h^2 h! \mu_{1,1}^{h-2} \rho_{1,2}\rho_{2,1} = \frac{R^- R^+}{m} \sum_{h=1}^{2g(n)} \frac{m^{h+1}}{(m)_{h+1}} h^2 (1+Q)^{h-2} \leq \frac{R^- R^+}{m} \int_0^{2g(n)} x^2 e^{\frac{xQ}{2}} dx \tag{4.21}$$

Where we eliminate the term $\frac{m^{k+1}}{(m)_{k+1}}$ in the above in the same way as in (4.19). An integral of the form seen in (4.21) can be evaluated by integrating by parts twice to deduce the

following (where $t > 0$),

$$\int_0^y x^2 e^{-tx} = \frac{2}{t^3} - e^{-ty}\left(\frac{2}{t^3} + \frac{2y}{t^2} + \frac{2y^2}{t}\right) \leq \frac{2}{t^3}.$$

Thus the expected number of these subgraphs in $\mathbb{DCM}$ can be bounded above by $\frac{16R^-R^+}{m|Q|^3} \to 0$.

The second case is $T(a,b)$ where we have a degree 4 vertex. In this case there are at most $hh!$ such subgraphs on h vertices ($\leq h$ choices of the two cycle lengths and assuming the associated automorphism groups are all trivial gives this bound). Again we apply Lemma 4.2.1 to compute the expected number of such subgraphs of size at most $2g(n)$ which this time is at most

$$\sum_{h=1}^{2g(n)} \frac{n^h}{(m)_{h+1}} h\mu_{1,1}^{h-1}\rho_{2,2} \leq \frac{R^+\Delta}{m}\sum_{h=1}^{2g(n)} \frac{m^{h+1}}{(m)_{h+1}} he^{-hQ} \leq \frac{R^+\Delta}{m}\int_0^{2g(n)} xe^{-\frac{xQ}{2}}dx \tag{4.22}$$

Note we may pick either $\rho_{2,2,} \leq \frac{m}{n}R^-\Delta$ or $\frac{m}{n}R^+\Delta$ here by selecting which part of the product $d_i^-(d_i^- - 1)d_i^+(d_i^+ - 1)$ to bound by Δ in computing $\rho_{2,2}$ and so we pick the smaller of the two. We may proceed similarly to before, integrating by parts which allows us to bound integrals of the form found in 4.22 as

$$\int_0^y x^2 e^{-tx} = \frac{1}{t^2} - e^{-ty}\left(\frac{1}{t^2} + \frac{y}{t}\right) \leq \frac{1}{t^2}.$$

So, we can bound (4.22) above by $\frac{4R^+\Delta}{m|Q|^2} \leq \frac{4(R^+R^-)^{2/3}}{m^{2/3}\varsigma^{1/3}|Q|^2}\frac{\Delta}{m^{1/3}} \to 0$. Thus by Markov's inequality there are no copies of $S(a,b,c)$ or $T(a,b)$ on at most $2g(n)$ vertices with high probability. Combining Lemma 4.4.1 and Lemma 4.4.4 we deduce there are no cycles of length at least $g(n)$ with high probability. As any complex strongly connected digraph contains a copy of at least one of these, we deduce that $\mathbb{DCM}$ contains no complex components with high probability.

4.4.4 Length of the kth largest cycle

Lemma 4.4.6. *The kth longest cycle in $\mathbb{DCM}$ follows the distribution from Theorem 4.1.3*

The idea now will be to apply a local coupling version of the Chen-Stein method (see [25, Theorem 2.8]) to deduce that the number of cycles in $\mathbb{DCM}$ of length between $\alpha/|Q|$ and $g(n)$ converges to a Poisson distribution of mean ξ_α. Let us start by stating the version of the Chen-Stein method which we will apply,

Theorem 4.4.7. Let $W = \sum_{i\in\Gamma} X_i$ be a sum of indicator variables and let $p_i := \mathbb{E}(X_i)$. For each $i \in \Gamma$, divide $\Gamma \setminus \{i\}$ into two sets, Γ_i^s and Γ_i^w. Define

$$Z_i := \sum_{j\in\Gamma_i^s} X_j \text{ and } W_i := \sum_{j\in\Gamma_i^w} X_j.$$

Suppose that there exist random variables, W_i^1 and $\widetilde{W}_i^1$ defined on the same probability space such that

$$\mathcal{L}(\widetilde{W}_i^1) = \mathcal{L}(W_i|X_i = 1) \text{ and } \mathcal{L}(W_i^1) = \mathcal{L}(W_i).$$

Then,

$$d_{\mathrm{TV}}(W, Po(\mathbb{E}(W))) \leq \min(1, \mathbb{E}(W)^{-1}) \sum_{i\in\Gamma} \Big(p_i\mathbb{E}(X_i + Z_i) + \mathbb{E}(X_iZ_i) + p_i\mathbb{E}|\widetilde{W}_i^1 - W_i^1|\Big) \tag{4.23}$$

Note that this lemma requires us to have a copy of $W_i|X_i = 1$. To create such a copy, we will use the following lemma to couple the configuration model with itself conditioned on the containment of a given subgraph.

Lemma 4.4.8. *Let $G = (A\cup B, E)$ be a balanced bipartite complete graph and let $\mathcal{M}$ be a uniformly chosen random perfect matching of G. Suppose that $a_1, a_2, \ldots, a_k$ are distinct elements of A and $b_1, b_2, \ldots, b_k$ are distinct elements of B. Then, the following procedure gives a copy of $\mathcal{M}|(a_1b_1, a_2b_2, \ldots, a_kb_k \in \mathcal{M})$:*

1. *Sample an element M from $\mathcal{M}$ and set $M_0 = M$.*

2. *For each $i \leq k$:*

   - *If $a_i b_i$ is an edge of M_{i-1} set $M_i = M_{i-1}$.*
   - *Otherwise, let a_i' be the unique neighbour of b_i and b_i' be the unique neighbour of a_i in M_{i-1}. Let $M_i = (M_{i-1} \setminus \{a_i b_i', a_i' b_i\}) \cup \{a_i b_i, a_i' b_i'\}$.*

That is M_k is sampled uniformly from $\mathcal{M}|(a_1 b_1, a_2 b_2, \ldots, a_k b_k \in \mathcal{M})$.

Proof. Note that it is sufficient to prove the lemma for $k = 1$ as if $\mathcal{B} = \mathcal{M}|(a_1 b_1, a_2 b_2, \ldots, a_{k-1} b_{k-1} \in \mathcal{M})$ then clearly, $\mathcal{B}|(a_k b_k \in \mathcal{B}) = \mathcal{M}|(a_1 b_1, a_2 b_2, \ldots, a_k b_k \in \mathcal{M})$. So, the general case follows by induction on the $k = 1$ case.

Now, we prove the lemma for $k = 1$. To do so, we will show that each of the $(n-1)!$ atoms of $\mathcal{M}|(a_1 b_1 \in \mathcal{M})$ comes from precisely n atoms of $\mathcal{M}$ via this switching approach. So let M be an atom of $\mathcal{M}|(a_1 b_1 \in \mathcal{M})$ and let $e = cd$ be an edge of M with $c \in A, d \in B$. Now, consider the inverse switching, $M \to (M \setminus \{ab, cd\}) \cup \{ad, bc\}$ (note if we chose the edge ab, then this is simply the identity $M \to M$). Thus, for each atom M of $\mathcal{M}|(a_1 b_1 \in \mathcal{M})$ there are n ways to get to an atom of $\mathcal{M}$. Similarly we can show that each atom of $\mathcal{M}$ is mapped to a unique element of $\mathcal{M}|(a_1 b_1 \in \mathcal{M})$. Thus, as $\mathcal{M}$ has the uniform distribution, so does the random variable obtained by our procedure above. □

Note that this lemma does not allow us to generate the configuration model conditioned on the existence of a subgraph specified in the usual way by the locations of its vertices unless all of the involved vertices have in- and out-degrees at most 1. This is due to the fact that Lemma 4.4.8 allows us to condition on which pairs of stubs are connected rather than which pairs of vertices. Instead we shall condition on the existence of *principal subgraphs* which we shall define as follows.

Definition 4.4.9. Let $\mathbb{DCM}$ be a configuration model random digraph with degree distribution (D_-, D_+). Let $M = M(\mathbb{DCM})$ be the associated perfect matching of in- and out-stubs. A *principal subgraph* of $\mathbb{DCM}$ is an event of the form $M' \subseteq M$ where M' is a partial matching of in- and out-stubs.

A similar notion was seen in [51] in which their canonical events are precisely the same as our principal subgraphs. Furthermore, note that the event corresponding to a subgraph in $\mathbb{DCM}$ may be written as a union of events corresponding to principal subgraphs. In particular, the existence of the cycle $v_1 v_2 \dots v_k$ in $\mathbb{DCM}$ can be written as the union of $\prod_{i=1}^{k} d^-_{\mathbb{DCM}}(v_i) d^+_{\mathbb{DCM}}(v_i)$ events corresponding to principal cycles. For this reason, in places where there may be some ambiguity as to whether we are dealing with principal subgraphs or not, we shall refer to subgraphs in the usual sense as *union subgraphs*.

This leaves us in a setting to which we can apply Theorem 4.4.7. We shall show that the number of cycles in $\mathbb{DCM}$ of lengths between $\alpha/|Q|$ and $g(n)$ is Poisson distributed with mean ξ_α. Let Γ be the set of all principal cycles which have lengths between $\alpha/|Q|$ and $g(n)$. For each $C \in \Gamma$ and digraph J on the same vertex set and stubs as $\mathbb{DCM}$ let $X_C(J)$ be the indicator function that C is a principal subgraph of J. Also, we split $\Gamma \setminus \{C\}$ into a set strongly dependent on C and a set weakly dependent on C. Define the strongly dependent set Γ^s_C to be the set of all principal cycles which share at least one vertex with C. The weakly dependent set Γ^w_C contains all of the other principal cycles, it is the set of principal cycles which are vertex disjoint from C. Finally, we define W^1_C to be W_C for an independent copy $\mathbb{DCM}'$ of the configuration model and $\widetilde{\mathbb{DCM}}_C$ to be obtained from $\mathbb{DCM}'$ by applying a 4-cycle switching to each edge of C in $\mathbb{DCM}_C$ in turn and define $\widetilde{W}^1_C$ in the obvious way to be W_C for the digraph $\widetilde{\mathbb{DCM}}_C$. That is,

$$W_C = \sum_{C' \in \Gamma^w_C} X_{C'}(\mathbb{DCM}') \qquad\qquad \widetilde{W}^1_C = \sum_{C' \in \Gamma^w_C} X_{C'}(\widetilde{\mathbb{DCM}}_C)$$

These variables clearly satisfy the assumptions of Theorem 4.4.7 and so we must bound the expectations in the statement of the theorem to compute an upper bound on $d_{\mathrm{TV}}(W, Po(\mathbb{E}(W)))$. The idea now is to reduce the whole problem to one of bounding the expected numbers of certain subgraphs being contained in $\mathbb{DCM}$, $\mathbb{DCM}'$ and $\widetilde{\mathbb{DCM}}_C$. For the remainder of this section, we write X_C for $X_C(\mathbb{DCM})$ and $p_C = \mathbb{E}(X_C)$ unless specified otherwise.

We shall bound the three terms from (4.23) one by one. First let us consider the term

$$\eta := \sum_{C\in\Gamma} p_C \mathbb{E}(X_C + Z_C) = \sum_{C\in\Gamma} \sum_{C'\in\Gamma^s_C\cup\{C\}} p_C p'_C \tag{4.24}$$

Note that the set, $\{C' \in \Gamma^s_C \cup \{C\}\}$ is the set of all cycles which share at least one vertex with C. Thus, the union of C and C' is a strongly connected digraph (here we allow multiple edges). In the computation of η we will use the following generalisation of Lemma 2.3.4. No changes are required to the proof of the lemma to generalise from digraphs to multi-digraphs.

Lemma 4.4.10. *Each strongly connected multi-digraph D with excess k may be formed in at most 27^k ways as the union of a pair of directed cycles C_1 and C_2.*

Define Ξ to be the set of all strongly connected multi-digraphs with vertices a subset of $V(\mathbb{DCM})$ such that all vertices have degrees $d^+(v) = d^-(v) = 1$ or $d^+(v) = d^-(v) = 2$ with at least one vertex which has $d^+(v) = d^-(v) = 2$. Note that Ξ is precisely the set of multi-digraphs which can be formed as the edge disjoint union of the two cycles $C \in \Gamma$ and $C' \in \Gamma^s_C$. Furthermore, note that the excess of a multi-digraph in Ξ is precisely the number of vertices which have $d^+(v) = d^-(v) = 2$. We let Ξ_k be the set $\Xi^h_k := \{F \in \Xi | |F| = h \text{ and excess}(F) = k\}$. Moreover, for each $F \in \Xi$ define

$$t(F) := \prod_{v\in V(F)} d^-_{\mathbb{DCM}}(v)^{d^-_F(v)} d^+_{\mathbb{DCM}}(v)^{d^+_F(v)} \tag{4.25}$$

Observe that for a given C, C' whose union is F, $t(F)$ is the number of pairs of principal cycles $\widetilde{C}, \widetilde{C}'$ which are copies of the same cycles as C, C'. Thus, combining (4.25) with Lemma 4.4.10 allows us to bound η as follows,

$$\eta = \sum_{C\in\Gamma} \sum_{C'\in\Gamma^s_C\cup\{C\}} p_C p_{C'} \leq \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{k=1}^{h} \sum_{F\in\Xi^h_k} \frac{27^k t(F)}{(m-h-k)^{h+k}} \tag{4.26}$$

Now, let Λ^h_k be the set of all strongly connected labelled multi-digraphs with h vertices

such that all vertices have degrees $d^+(v) = d^-(v) = 1$ or $d^+(v) = d^-(v) = 2$ with precisely k vertices v such that $d^+(v) = d^-(v) = 2$. This allows us to write

$$\sum_{F \in \Xi_k^h} t(F) = \frac{1}{h!} \sum_{F \in \Lambda_k^h} \sum_{\phi : V(F) \hookrightarrow V(\mathbb{DCM})} \prod_{v \in V(F)} d^-_{\mathbb{DCM}}(v)^{d^-_F(v)} d^+_{\mathbb{DCM}}(v)^{d^+_F(v)} \tag{4.27}$$

Where the division by $h!$ comes from the fact that Λ_k^h distinguishes between digraphs obtained from one-another by a permutation of the vertex set. Arguing similarly to Lemma 4.2.1, noting that we know the degree sequence of F and using $\mu_{1,1} = \frac{m}{n}(1+Q) \leq \frac{m}{n}$, we deduce that

$$\sum_{\phi : V(F) \hookrightarrow V(\mathbb{DCM})} \prod_{v \in V(F)} d^-_{\mathbb{DCM}}(v)^{d^-_F(v)} d^+_{\mathbb{DCM}}(v)^{d^+_F(v)} \leq n^h \mu_{1,1}^{h-k} \mu_{2,2}^k \leq m^h \mu_{2,2}^k \mu_{1,1}^{-k}.$$

Substituting into (4.27) and applying Lemma 4.3.1 we find

$$\begin{aligned} \sum_{F \in \Xi_k^h} t(F) &\leq \frac{k}{h+k} \frac{(h+k)!}{h!} \binom{h}{k} m^h \mu_{2,2}^k \mu_{1,1}^{-k} = \frac{k}{h+k} \binom{h+k}{2k} \frac{(2k)!}{k!} m^h \mu_{2,2}^k \mu_{1,1}^{-k} \\ &\leq \frac{(h+k)^{2k-1}}{(k-1)!} m^h \mu_{2,2}^k \mu_{1,1}^{-k}. \end{aligned} \tag{4.28}$$

To finish, we note that $((m-h-k)^{h+k}) = (1+o(1))m^{h+k}$ as $h+k = o(m^{1/2})$, this allows us to substitute the bound found in (4.28) into (4.26) where we deduce

$$\begin{aligned} \eta &\leq (1+o(1)) \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{k=1}^{h} \frac{27^k}{m^{h+k}} \frac{(h+k)^{2k-1}}{(k-1)!} m^h \mu_{2,2}^k \mu_{1,1}^{-k} \\ &\leq (1+o(1)) \frac{216 g(n) \mu_{2,2}}{m} \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{k=0}^{h} \frac{108^k h^{2k} \mu_{2,2}^k}{m^k k! \mu_{1,1}^k}. \end{aligned} \tag{4.29}$$

Part of the expression in (4.29) is in the form of an exponential sum. Evaluating this

sum allows us to deduce that $\eta = o(1)$.

$$\eta \leq (1+o(1))\frac{216g(n)\mu_{2,2}}{m\mu_{1,1}} \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} e^{\frac{108h^2\mu_{2,2}}{m\mu_{1,1}}} \leq (1+o(1))\frac{432g(n)^2\mu_{2,2}}{m\mu_{1,1}} e^{\frac{432g(n)^2\mu_{2,2}}{m\mu_{1,1}}} = o(1). \tag{4.30}$$

Where we deduce that this is $o(1)$ in the same way as when we bound (4.22).

Next we shall consider the term,

$$\theta := \sum_{C\in\Gamma} \mathbb{E}(X_C Z_C) = \sum_{C\in\Gamma} \sum_{C'\in\Gamma^s_C\cup\{C\}} \mathbb{E}(X_C X'_C) \tag{4.31}$$

So note that $\mathbb{E}(X_C X_{C'})$ is the probability that both C and C' are simultaneously present. Furthermore, note that $C \cup C'$ is a strongly connected (not necessarily simple) digraph with maximum degree at most 4. Thus we can use a similar strategy to the one used to bound η. So define Θ to be the set of all strongly connected multi-digraphs F with $V(F) \subseteq V(G)$ and $\Delta(F) \leq 4$. Furthermore for $a, b \in \mathbb{N}$, define

$$\Theta^h_{a,b} := \{F \in \Theta | |F| = h, n_{1,2}(F) = n_{2,1}(F) = a, n_{2,2}(F) = b\}$$

Where we define $\Theta^h_{0,0} = \emptyset$. Define $t(F)$ for $F \in \Theta$ in the same way as (4.25). Then, $t(F)$ for a given construction, $F = C \cup C'$ from a pair of principal cycles, $t(F)$ is again the number of pairs of principal cycles $\widetilde{C}, \widetilde{C}'$ which are copies of the same cycles as C, C'. Thus, if we apply Lemma 4.4.10 we get the following bound on θ,

$$\theta = \sum_{C\in\Gamma} \sum_{C'\in\Gamma^s_C} \mathbb{E}(X_C X_{C'}) \leq \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{a=0}^{h} \sum_{b=0}^{h} \sum_{F\in\Theta^h_{a,b}} \frac{27^k t(F)}{(m-h-a-b)^{h+a+b}} \tag{4.32}$$

Now, we define $\Lambda^h_{a,b}$ to be the set of all strongly connected labelled multi-digraphs F with h vertices such that $n_{1,1}(F) = h - 2a - b$, $n_{1,2}(F) = n_{2,1}(F) = a$ and $n_{2,2}(F) = b$. (Note

$\Lambda^h_{0,k} = \Lambda^h_k$). Thus, arguing as previously,

$$\sum_{F\in\Theta^h_{a,b}} t(F) = \frac{1}{h!}\sum_{F\in\Lambda^h_{a,b}}\sum_{\phi:V(F)\hookrightarrow V(\mathbb{DCM})}\prod_{v\in V(F)} d^-_{\mathbb{DCM}}(v)^{d^-_F(v)} d^+_{\mathbb{DCM}}(v)^{d^+_F(v)} \tag{4.33}$$

Again, we may argue similarly to Lemma 4.2.1 to deduce that for $F \in \Lambda^h_{a,b}$,

$$\sum_{\phi:V(F)\hookrightarrow V(\mathbb{DCM})}\prod_{v\in V(F)} d^-_{\mathbb{DCM}}(v)^{d^-_F(v)} d^+_{\mathbb{DCM}}(v)^{d^+_F(v)} \leq m^h \mu^a_{1,2}\mu^a_{2,1}\mu^b_{2,2}\mu^{-(2a+b)}_{1,1}.$$

We can substitute this into (4.32) and apply Lemma 4.3.1 to find

$$\begin{aligned}\sum_{F\in\Theta^h_{a,b}} t(F) &\leq \frac{3a+2b}{h+a+b}\frac{(h+a+b)!}{h!}\binom{h}{a,a,b} m^h \mu^a_{1,2}\mu^a_{2,1}\mu^b_{2,2}\mu^{-(2a+b)}_{1,1} \\ &\leq \frac{3a+2b}{h} 2^{a+b}\frac{h^{3a+2b}}{a!a!b!} m^h \mu^a_{1,2}\mu^a_{2,1}\mu^b_{2,2}\mu^{-(2a+b)}_{1,1} \\ &= \frac{3a+2b}{h}\frac{1}{a!a!b!}\left(\frac{2h^3\mu_{1,2}\mu_{2,1}}{\mu^2_{1,1}}\right)^a\left(\frac{2h^2\mu_{2,2}}{\mu_{1,1}}\right)^b m^h \end{aligned} \tag{4.34}$$

Note that this bound is a sum of two terms due to the factor $3a + 2b$ in (4.34). So we will split this bound into $t_a(F) + t_b(F)$ in the obvious way. This allows us to bound $\theta \leq \theta_a + \theta_b$ by only considering the $t_a(F)$ or $t_b(F)$ terms which will be convenient for us. We substitute the bound from (4.34) into (4.32) and use that $(m - h - a - b)^{h+a+b} = (1 + o(1))m^{h+a+b}$ to deduce

$$\theta_a \leq (1+o(1))\sum_{h=\frac{\alpha}{|Q|}}^{2g(n)}\sum_{a=0}^{h}\sum_{b=0}^{h}\frac{3a}{h}\frac{1}{a!a!b!}\left(\frac{2\cdot 27^2 h^3\mu_{1,2}\mu_{2,1}}{m\mu^2_{1,1}}\right)^a\left(\frac{54h^2\mu_{2,2}}{m\mu_{1,1}}\right)^b \tag{4.35}$$

$$\theta_b \leq (1+o(1))\sum_{h=\frac{\alpha}{|Q|}}^{2g(n)}\sum_{a=0}^{h}\sum_{b=0}^{h}\frac{2b}{h}\frac{1}{a!a!b!}\left(\frac{2\cdot 27^2 h^3\mu_{1,2}\mu_{2,1}}{m\mu^2_{1,1}}\right)^a\left(\frac{54h^2\mu_{2,2}}{m\mu_{1,1}}\right)^b \tag{4.36}$$

Note that in both cases we can evaluate the two inner sums in terms of exponential functions and modified Bessel functions of the first kind. However the latter of these is a little difficult to work with, hence we will replace the $a!a!$ with $(2a)!$ allowing us to use

hyperbolic trigonometric functions instead. In particular, we have

$$\frac{1}{a!a!} \leq \frac{4^a}{(2a)!} \quad \text{and} \quad \frac{1}{a!(a-1)!} \leq \frac{4^a}{(2a-1)!}.$$

Looking first at θ_a we get the bound

$$\begin{aligned}
\theta_a &\leq (1+o(1)) \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{a=0}^{h} \sum_{b=0}^{h} \frac{3}{h} \frac{1}{(2a-1)!b!} \left(\frac{8 \cdot 27^2 h^3 \mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2} \right)^a \left(\frac{54h^2\mu_{2,2}}{m\mu_{1,1}} \right)^b \\
&= (1+o(1))162\sqrt{2} \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sqrt{\frac{h\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \sinh\left(54\sqrt{\frac{2h^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \right) e^{\frac{54h^2\mu_{2,2}}{m\mu_{1,1}}} \\
&\leq (1+o(1))648\sqrt{\frac{g(n)^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \sinh\left(216\sqrt{\frac{g(n)^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \right) e^{\frac{216g(n)^2\mu_{2,2}}{m\mu_{1,1}}}. \qquad (4.37)
\end{aligned}$$

Where the final inequality follows from the fact that $\sqrt{x}$, $\sinh(x)$ and e^x are all increasing for $x > 0$. Bounding θ_b is similar,

$$\begin{aligned}
\theta_b &\leq (1+o(1)) \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \sum_{a=0}^{h} \sum_{b=1}^{h} \frac{2}{h} \frac{1}{(2a)!(b-1)!} \left(\frac{8 \cdot 27^2 h^3 \mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2} \right)^a \left(\frac{54h^2\mu_{2,2}}{m\mu_{1,1}} \right)^b \\
&= (1+o(1))108 \sum_{h=\frac{\alpha}{|Q|}}^{2g(n)} \frac{h\mu_{2,2}}{m\mu_{1,1}} \cosh\left(54\sqrt{\frac{2h^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \right) e^{\frac{54h^2\mu_{2,2}}{m\mu_{1,1}}} \\
&\leq (1+o(1))432\frac{g(n)^2\mu_{2,2}}{m\mu_{1,1}} \cosh\left(216\sqrt{\frac{g(n)^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2}} \right) e^{\frac{216g(n)^2\mu_{2,2}}{m\mu_{1,1}}} \qquad (4.38)
\end{aligned}$$

Both (4.37) and (4.38) are $o(1)$ which follows from the facts that

$$\frac{g(n)^2\mu_{2,2}}{m\mu_{1,1}} \to 0 \qquad \text{and} \qquad \frac{g(n)^3\mu_{1,2}\mu_{2,1}}{m\mu_{1,1}^2} \to 0.$$

The first of which we showed in (4.30) and the latter follows directly from the assumption that $g(n)^3 = o(\frac{m}{R^- R^+})$. Hence $\theta = o(1)$. Finally, we will bound the terms,

$$\kappa_C := \mathbb{E}|\widetilde{W}_C^1 - W_C^1|. \qquad (4.39)$$

We note that unlike in bounding η and θ we will not need to look at the sum over $C \in \Gamma$ in computing the bound. This is due to the fact that κ_C depends on a much more global structure than the terms $p_C\mathbb{E}(X_C + Z_C)$ and $\mathbb{E}(X_C Z_C)$. For the bounding of κ_C, the first thing to observe is that due to the choice of coupling and Γ_C^s we have $\widetilde{W}_C^1 \geq W_C^1$. This is because all of the edges which are removed by the switchings are incident to a vertex of C and therefore none of the principal cycles containing these edges are in Γ_C^w which are the cycles contributing to $\widetilde{W}_C^1 \geq W_C^1$. This enables us to bound κ_C by computing the number of expected cycles from Γ_C^w which are added by the switchings.

In order to add a cycle C' with the switchings over 4-cycles which produce $\widetilde{\mathbb{DCM}}_C$ from $\mathbb{DCM}$ the following must be true for some k,

- All but k edges of C' must be present in $\mathbb{DCM}$.
- k of the edges of C must add these edges after applying the switching.

Now let us compute the probability that this event comes to pass. So let us fix k edges missing from C' which we match up with k of the edges of C which will be used to add them when we apply the switching. So if one such edge is $e = uv$ and this is matched with $e' = u'v'$, in order for the switching to add e, before switching we must have edges uv' and $u'v$ present in G (note that because the edges are directed we do not need to consider the alternate switching using uu' and vv' like we would if working with graphs). Hence in total there are $|C'| + k$ edges which we must find in $\mathbb{DCM}$ before switching in order that C is a cycle of $\widetilde{\mathbb{DCM}}_C$. The probability we find these edges is $((m)_{|C'|+k})^{-1}$.

Now, let us count how many such structures there are which produce the same union cycle as C'. As C is a principal cycle, we know exactly which stubs we use for it and so there is no contribution to the number of copies from vertices of C. However there are $d^+(u)d^-(v)$ switchings which add the union edge uv and so the number of ways in which the switching structure with a union cycle which is that of C' can be found is

$$\prod_{v \in V(C')} d^-(v)d^+(v).$$

Thus for any fixed set of missing edges for the union cycle with the same edges as C' and choice of edges from C with which we add these edges, the expected number of additional principal cycles after the switchings is

$$\frac{1}{(m)_{|C'|+k}} \prod_{v \in V(C')} d^-(v)d^+(v). \tag{4.40}$$

There are $\binom{|C|}{k}$ ways to pick the edges of C which we switch over to add the missing edges of C'. Also there are $\binom{|C'|}{k}$ ways to pick the missing edges of C' and $k!$ ways to assign edges of C to missing edges of C'. Thus, when we sum over principal cycles C', missing edges of C and C' and matchings of the missing edges, we find the expected number of such structures is at most

$$\mathbb{E}|\widetilde{W}_C^1 - W_C^1| \leq \sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \sum_{k=1}^{\min(|C|,h)} \frac{1}{h} \sum_{C' \subseteq G; |C'|=h} \binom{|C|}{k}\binom{h}{k} \frac{k!}{(m)_{h+k}} \prod_{v \in V(C')} d^-(v)d^+(v). \tag{4.41}$$

Applying standard bounds on binomial coefficients and falling factorials and taking vertices from $V(G)$ for $|C'|$ with replacement rather than without allows us to bound (4.41) as follows

$$\begin{aligned}
\mathbb{E}|\widetilde{W}_C^1 - W_C^1| &\leq (1+o(1)) \sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \sum_{k=1}^{\infty} \frac{|C|^k h^{k-1}}{k! m^k} (1+Q)^h \\
&\leq (1+o(1)) \sum_{|C'|=\frac{\alpha}{|Q|}}^{g(n)} \frac{e^{\frac{|C|h}{m}} - 1}{|C'|} \\
&\leq \frac{g(n)|Q|}{\alpha} \left(e^{\frac{g(n)^2}{m}} - 1 \right) \leq \frac{2g(n)^3|Q|}{\alpha m} = o(1).
\end{aligned} \tag{4.42}$$

Where we note that $g(n)^2/m < 1$ allows us to use the bound $e^x - 1 \leq 2x$ in (4.42). So $\kappa_C \leq \kappa = o(1)$ for all C where $\kappa = \frac{2g(n)^3|Q|}{\alpha m}$. Finally, note that this implies $\sum_{C \in \Gamma} p_C \kappa_C \leq \kappa \mathbb{E}(W)$. We subsequently show that $\mathbb{E}(W) \leq \xi_\alpha + o(1)$ from which it follows that $\sum_{C \in \Gamma} p_C \kappa_C = o(1)$ as ξ_α is a constant independent of n. Using this and the

fact that $\eta, \theta = o(1)$, we conclude that

$$d_{\mathrm{TV}}(W, Po(\mathbb{E}(W))) = o(1).$$

Note that W is the number of cycles of $\mathbb{DCM}$ of length between $\alpha/|Q|$ and $g(n)$. By Lemma 4.4.1 and Lemma 4.4.4, the probability that there are any longer cycles is $o(1)$. Thus, if W' is the number of cycles of $\mathbb{DCM}$ of length at least $\alpha/|Q|$, $d_{\mathrm{TV}}(W', Po(\mathbb{E}(W))) = d_{\mathrm{TV}}(W, Po(\mathbb{E}(W))) + o(1) = o(1)$. To show that the mean of the corresponding Poisson distribution can be taken to be ξ_α note that $d_{\mathrm{TV}}(Po(\lambda), Po(\mu)) \leq |\lambda - \mu|$ holds for all $\lambda, \mu \geq 0$. This follows from coupling $\mathrm{Bin}(n, \lambda/n)$ and $\mathrm{Bin}(n, \mu/n)$ by coupling their constituent Bernoulli trials and noting that $d_{\mathrm{TV}}(\mathrm{Bin}(n, \lambda/n), Po(\lambda)) = o_n(1)$. Thus it suffices to show that $\mathbb{E}(W) = \xi_\alpha + o(1)$. Note that W is the number of cycles in $\mathbb{DCM}$ of lengths between $\alpha/|Q|$ and $g(n)$. Thus upper bounding $\mathbb{E}(W)$ may be done in the same way as the proof of Lemma 4.4.4. However, we can be more careful than we are in (4.20) as in this case $2h^2/m = o(1)$, so this line can be replaced by

$$\sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \frac{e^{hQ+\frac{2h^2}{m}}}{h} \leq \sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \frac{e^{hQ+o(1)}}{h} \leq (1+o(1)) \int_{\frac{\alpha}{|Q|}}^{\infty} \frac{e^{hQ}}{h} dh = (1+o(1)) \int_{\alpha}^{\infty} \frac{e^{-\lambda}}{\lambda} d\lambda = \xi_\alpha + o(1). \tag{4.43}$$

Next, we lower bound $\mathbb{E}(W)$ by $\xi_\alpha - o(1)$. In order to do this we will use Lemma 4.2.2 in combination with the inequality

$$1 - x \geq e^{-x-x^2} \quad \text{for} \quad x \leq \frac{1}{2}.$$

This allows us to deduce that the probability of finding a cycle of length between $\alpha/|Q|$ and $g(n)$ is at least

$$\sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \frac{h!}{h} \binom{n}{h} \frac{(1+Q)^h}{(n)_h} \left(1 - \frac{2h^2\Delta^2}{\varepsilon n}\right) \left(1 - \frac{h\Delta^2}{2m}\right) \geq \sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \frac{(1+o(1))}{h} e^{h(Q-\frac{Q^2}{2})} \tag{4.44}$$

Now, that the exponent in the above is equal to $hQ - o(1)$ follows from the assumption $g(n) = o(1/|Q|^2)$. Thus in analogy with (4.43), we deduce that (4.44) is at most

$$(1+o(1)) \sum_{h=\frac{\alpha}{|Q|}}^{g(n)} \frac{e^{hQ-o(1)}}{h} \geq (1-o(1)) \int_{\frac{\alpha}{|Q|}}^{\infty} \frac{e^{hQ}}{h} dh = (1-o(1)) \int_{\alpha}^{\infty} \frac{e^{-\lambda}}{\lambda} d\lambda = \xi_\alpha - o(1). \tag{4.45}$$

Combining (4.43) and (4.45) we deduce that $\mathbb{E}(W) = \xi_\alpha + o(1)$ as required.

4.4.5 Putting it all together

We can deduce the statement of the theorem from the previous sections as follows. First we apply Lemma 4.4.1, Lemma 4.4.4 and Lemma 4.4.5 to deduce there are no cycles with length at least $\omega(1/|Q|)$ or complex components. Then, we can apply Lemma 4.4.6 to deduce that the distribution of the number of cycles with at least $\alpha/|Q|$ vertices converges to a Poisson distribution with mean ξ_α from which the statement that

$$\mathbb{P}\left(|\mathcal{C}_k| \geq \frac{\alpha}{|Q|}\right) = 1 - \sum_{i=0}^{k-1} \frac{\xi_\alpha^i}{i!} e^{-\xi_\alpha} + o(1)$$

follows immediately as the above is simply the probability that a Poisson(ξ_α) distribution is at least k (plus the $o(1)$ term).

4.5 Concluding Remarks

In this chapter, we showed that the largest component of the directed configuration model is of order $|Q|^{-1}$ when $nQ^3(R^-R^+) \to -\infty$ and found the distribution of the size of the kth largest component for any k. In a subsequent work [14] we shall show that under similar conditions to those in this chapter that for degree sequences such that $nQ^3(R^-R^+)^{-1} \to \infty$ the largest component is of order $nQ^2(R^-R^+)^{-1}$. This quantity matches the $1/|Q|$ we find in this chapter if $n|Q|^3(R^-R^+)^{-1} \to c$ for some constant c and suggests that one may find a critical window phenomenon for degree sequences with

such parameters. Note that in particular this is precisely the critical window of $D(n,p)$ (looking at typical degree sequences from the model $D(n,p)$). Moreover, the recent result of Donderwinkel and Xie certainly seems to indicate that the point $Q = 0$ lies inside a critical window.

An interesting question for further study would be to ask about the joint distribution of the largest strongly connected components in the directed configuration model with parameters as in this chapter. The fact that we find

$$\mathbb{P}\left(|\mathcal{C}_k| \geq \frac{\alpha}{|Q|}\right) = 1 - \sum_{i=0}^{k-1} \frac{\xi_\alpha^i}{i!} e^{-\xi_\alpha} + o(1)$$

seems rather suggestive of an underlying Poisson process. As such we make the following conjecture.

Conjecture 4.5.1. Let $\mathbb{DCM}$ be a directed configuration model with parameters as in Theorem 4.1.3 and suppose the sizes of its components in descending order are given by the random variables $Z_1 \geq Z_2 \geq \ldots$. Suppose further that $X_1 \leq X_2 \leq \ldots$ are the points of a Poisson process of rate 1 and $Y_1 \geq Y_2 \geq \ldots$ are the unique positive solutions to

$$X_i = \int_{Y_i}^{\infty} \frac{e^{-x}}{x} dx.$$

Then we have that

$$(|Q|Z_1, |Q|Z_2, \ldots) \to (Y_1, Y_2, \ldots) \text{ as } n \to \infty.$$

Note that the above is also an open question for $D(n,p)$ with $p = (1-\varepsilon)/n$ and $\varepsilon \to 0$, $\varepsilon^3 n \to \infty$.

CHAPTER 5

WEAK COMPONENTS OF THE DIRECTED CONFIGURATION MODEL

5.1 Introduction

The study of the component structure of random graphs with given degrees, and in particular of the configuration model (CM), was pioneered by the work of Molloy and Reed [58], who provided a criterion to determine if a degree sequence typically produces a linear order connected component (known as the *giant*) or its largest component has sublinear order. Since then, it has become one of the central topics in random graph theory [8, 28, 35, 43, 45, 59].

Directed models are much less understood. A strongly connected component (SCC) of a directed graph is a maximal sub-digraph in which there exists a directed path between any ordered pair of nodes. Newman, Strogatz and Watts [64] initiated the study of the directed configuration model (DCM), and located the threshold for the existence of a giant SCC. Later, Cooper and Frieze [13] provided a rigorous proof for the existence of such threshold under certain conditions of the degree sequence. This problem has been recently revisited by the first and the third author [11], extending the range of applicability of the result. The component structure of the directed Erdős Rényi model and of other random directed graphs has been extensively studied [12, 15, 47, 55].

A weakly connected component (WCC) of a directed graph is a maximal sub-digraph

in which there exists an path from every node to every other node. WCCs naturally arise in areas such as epidemiology [26], data mining [48] or communication networks [10]. In the physics community, the study of weakly connected components has been neglected under the assumption that it effectively behaves like the undirected case (see, e.g., [64]). Kryven [49] observed that this assumption is wrong and predicted an alternative threshold for the appearance of the giant WCC, supported with an analytical but non-rigorous approach based on generating functions for bounded bi-degree distributions. The aim of this chapter is to provide a formal proof for the existence of the giant WCC threshold in the directed configuration model under the much weaker assumption of bounded second moments.

Let $[n] := \{1, \dots, n\}$ be a set of n vertices. Let $\vec{\mathbf{d}}_n = ((d_1^-, d_1^+), \dots, (d_n^-, d_n^+))$ be a bi-degree sequence with $m_n := \sum_{i\in[n]} d_i^+ = \sum_{i\in[n]} d_i^-$. Let $\Delta_n := \max_{i\in[n]}\{d_i^-, d_i^+\}$. The directed configuration model (DCM), $\mathbb{DCM} = \mathbb{DCM}(\vec{\mathbf{d}}_n)$, is the random directed multigraph on vertex set $[n]$ generated by assigning d_i^- in half-edges (*heads*) and d_i^+ out half-edges (*tails*) to vertex i, and then choosing a uniformly random matching between the set of heads and the set of tails. Let $n_{k,\ell} := \{i : (d_i^-, d_i^+) = (k, \ell)\}$. Let $D_n = (D_n^-, D_n^+)$ be the degree pair of a vertex chosen uniformly at random, that is $\mathbb{P}(D_n = (k, \ell)) = n_{k,\ell}/n$.

Let $(\vec{\mathbf{d}}_n)_{n\geq 1}$ be a sequence of bi-degree sequences. Unless specified otherwise, we will consider sequences that satisfy the following:

Condition 5.1.1. There exists a discrete probability distribution $D = (D^-, D^+)$ on $\mathbb{Z}_{\geq 0}^2$ with $\lambda_{k,\ell} := \mathbb{P}(D = (k, \ell))$ such that we have:

(i) convergence in distribution:

$$\lim_{n\to\infty} \frac{n_{k,\ell}}{n} = \lambda_{k,\ell}, \quad \text{for every } k, \ell \in \mathbb{Z}_{\geq 0}; \tag{5.1}$$

(ii) convergence of expected values:

$$\lim_{n\to\infty} \mathbb{E}[D_n^-] = \lim_{n\to\infty} \mathbb{E}[D_n^+] = \mathbb{E}[D^-] = \mathbb{E}[D^+] =: \lambda \in (0, \infty); \tag{5.2}$$

(iii) convergence of second moments: letting $(x)_a := x(x-1)\dots(x-a+1)$, for all $i, j \in \mathbb{N}$ and $i + j = 2$,

$$\lim_{n\to\infty} \mathbb{E}[(D_n^-)_i (D_n^+)_j] = \mathbb{E}[(D^-)_i (D^+)_j] =: \mu_{i,j} \in (0, \infty), \tag{5.3}$$

Define the *in- and out-size biased* distributions of D, denoted by D_{in} and D_{out} respectively, by

$$\mathbb{P}\left(D_{\text{in}} = (k-1, \ell)\right) = \frac{k\lambda_{k,\ell}}{\lambda}, \qquad \mathbb{P}\left(D_{\text{out}} = (k, \ell-1)\right) = \frac{\ell\lambda_{k,\ell}}{\lambda}. \tag{5.4}$$

Consider the random matrix

$$\Xi = \begin{pmatrix} D_{\text{out}}^- & D_{\text{out}}^+ \\ D_{\text{in}}^- & D_{\text{in}}^+ \end{pmatrix}, \tag{5.5}$$

where D_{out} and D_{in} are independent. Condition 5.1.1 implies that Ξ has a finite mean matrix

$$M := \frac{1}{\lambda} \begin{pmatrix} \mu_{1,1} & \mu_{0,2} \\ \mu_{2,0} & \mu_{1,1} \end{pmatrix}, \tag{5.6}$$

with largest eigenvalue

$$\rho := \frac{1}{\lambda} \left(\mu_{1,1} + \sqrt{\mu_{2,0}\mu_{0,2}}\right). \tag{5.7}$$

Let $\mathbf{q} = (q_-, q_+)$ be the extinction probability vector of a 2-type branching process with offspring Ξ.

Let $\mathcal{W}_n$ be the largest WCC in $\mathbb{DCM}$. (If there is more than one such WCC, we choose an arbitrary one among them as $\mathcal{W}_n$.) Let $v(\mathcal{W}_n)$ and $e(\mathcal{W}_n)$ be the number of vertices and edges in $\mathcal{W}_n$, respectively. Our main result is that the existence of a giant WCC undergoes a phase transition at $\rho = 1$:

Theorem 5.1.2. *Suppose that $(\vec{\mathbf{d}}_n)_{n\geq 1}$ satisfies Condition 5.1.1. If $\rho > 1$, then*

$$\frac{v(\mathcal{W}_n)}{n} \to \eta \in (0,1], \quad \textit{and} \quad \frac{e(\mathcal{W}_n)}{n} \to \zeta \in (0,\lambda], \tag{5.8}$$

in probability, where

$$\eta := \sum_{k,\ell\geq 0} \lambda_{k,\ell}(1-q_-^k q_+^\ell), \quad \textit{and} \quad \zeta := \sum_{k,\ell\geq 0} k\lambda_{k,\ell}(1-q_-^k q_+^\ell) = \sum_{k,\ell\geq 0} \ell\lambda_{k,\ell}(1-q_-^k q_+^\ell). \tag{5.9}$$

If $\rho < 1$, then

$$\frac{v(\mathcal{W}_n)}{n} \to 0, \quad \textit{and} \quad \frac{e(\mathcal{W}_n)}{n} \to 0, \tag{5.10}$$

in probability.

Remark 5.1.3 (Comparison with strongly connected components)**.** The SCC's of $\mathbb{DCM}$ have been studied in [11, 13] under Condition 5.1.1. Denote by $\mathcal{S}_n$ the largest SCC. Then, if $\mu_{1,1} > \lambda$,

$$\frac{v(\mathcal{S}_n)}{n} \to \eta_{\text{SCC}} := \sum_{k,\ell\geq 0} \lambda_{k,\ell}(1-r_-^k)(1-r_+^\ell), \tag{5.11}$$

where r_- and r_+ are extinction probabilities of branching processes with offspring distributions D_{out}^- and D_{in}^+, respectively. If $\mu_{1,1} < \lambda$, then

$$\frac{v(\mathcal{S}_n)}{a(n)} \to 0, \tag{5.12}$$

for any $a(n) \to \infty$ as $n \to \infty$. So, the existence of a giant SCC undergoes a phase transition at $\mu_{1,1} = \lambda$.

Our results combined with (5.11) and (5.12) indicate that the "separation" between the thresholds for the appearance of a WCC and a SCC giant in $\mathbb{DCM}$ is solely determined by the second moments of the marginals of the bi-degree distribution, and independent from the correlation between in- and out-degrees. A formal way to state it is through bond percolation. For any $p = p(n) \in [0,1]$, denote by $\mathbb{DCM}_p(\vec{\mathbf{d}}_n)$ the p-percolated directed configuration model obtained by sampling $\mathbb{DCM}(\vec{\mathbf{d}}_n)$ and independently retain

every directed edge with probability p. Given a degree sequence $\vec{\mathbf{d}}_n$, we define its p-thinned version $\vec{\mathbf{d}}_n^p$ as the random degree sequence where the degree D_n^p of a uniformly random vertex satisfies

$$\mathbb{P}\left(D_n^p = (k,\ell)\right) = \sum_{\substack{a\geq k\\ b\geq \ell}} \binom{a}{k}\binom{b}{\ell} p^{k+\ell}(1-p)^{a+b-(k+\ell)}\mathbb{P}\left(D_n = (a,b)\right) \ .$$

It is easy to show that $\mathbb{DCM}_p(\vec{\mathbf{d}}_n)$ is distributed in law as $\mathbb{DCM}(\vec{\mathbf{d}}_n^p)$, where $\vec{\mathbf{d}}_n^p$ is conditioned on having the same number of heads and tails (for an undirected analogue of this result, see e.g. Lemma 3.2 in [27]).

Define $p_{\text{SCC}} = \sqrt{\frac{\lambda}{\mu_{1,1}}}$ and $p_{\text{WCC}} = \sqrt{1/\rho} = \sqrt{\frac{\lambda}{\mu_{1,1}+\sqrt{\mu_{2,0}\mu_{0,2}}}}$, and note that $p_{\text{WCC}} < p_{\text{SCC}}$. Combining the previous results, we have the following two-point threshold phenomenon:

- if $p \in [0, p_{\text{WCC}})$: with high probability (whp) no a giant WCC exists.
- if $p \in (p_{\text{WCC}}, p_{\text{SCC}})$: whp a giant WCC, but no giant SCC exists;
- if $p \in (p_{\text{SCC}}, 1]$: whp a giant SCC exists;

So, for all degree sequences, there is a non-trivial regime where typically we see a giant WCC but no giant SCC. If $\vec{\mathbf{d}}_n$ satisfies $\mu_{1,1} = \lambda$, then p_{WCC} does not depend on the in-/out-degree correlation, formalising the separation intuition given above.

Remark 5.1.4 (The Community Configuration Model)**.** Motivated by the presence of clustering in real-world complex networks, recently the Borgs et al. have introduced the community configuration model (CCM) [9, 68], extending the Coloured Configuration Model of Kryven [4, 50]. In the CCM, each half-edge is assigned a colour in $[k]$, and a permutation matrix of size $k \times k$ provides the rules for matching the colours. The $\mathbb{DCM}$ can be seen as a particular case of the $\mathbb{CCM}$ with $k = 2$ and permutation matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In an upcoming work, we will study the component structure of the CCM in the general setting.

Remark 5.1.5 (Conditions on degree sequences)**.** Theorem 5.1.2 fails for sequences with infinite second moment. For instance, let $\Delta = n^{1/2+\epsilon}$ for $\epsilon > 0$ and consider the degree sequence with one vertex with degrees $(\Delta, 0)$ and all other vertices with degrees $(1, 0)$ or $(0, 1)$. Then the largest eigenvalue of the mean matrix M_n defined in (5.51) satisfies $\rho_n \to \infty$ but clearly the largest WCC has order $O(n^{1/2+\epsilon})$.

Under Condition 5.1.1, Theorem 5.1.2 can be transferred to simple random digraphs [5]. If the second moments diverge, then the probability that the $\mathbb{DCM}$ is simple goes to 0 as $n \to \infty$ and a different approach is needed to obtain results in the simple setting.

5.2 Multitype branching processes

Fix $p \in \mathbb{N}$. We write $\mathbf{0}$ and $\mathbf{1}$ for the all zeros and all ones row vectors of length p, respectively. Let $\mathbf{e}_a$ be $\mathbf{0}$ with the a-th component changed to 1. For any $\omega \in \mathbb{R}$, we write $\boldsymbol{\omega} = \omega\mathbf{1}$. For any two vectors $\mathbf{x}, \mathbf{y}$, we write $\mathbf{x} < \mathbf{y}$ (or $\mathbf{x} \leq \mathbf{y}$) to denote that $\mathbf{x}$ is less (or less or equal) than $\mathbf{y}$ in each component, and $\mathbf{x} \not< \mathbf{y}$ (or $\mathbf{x} \not\leq \mathbf{y}$) otherwise. Let $\Xi = (\xi_{ij})$ be a random $p \times p$ matrix with entries in $\mathbb{Z}_{\geq 0}$ in which rows are independent. Let $(\Xi(m;t)))_{m\geq 1, t\geq 0}$ be independent and identically distributed (iid) copies of Ξ. Let $\mathbf{z} = (z_1, \ldots, z_p)$. For $i \in [p]$, define the generating function

$$h_i(\mathbf{z}) = \sum_{k_1,\ldots,k_p \geq 0} \mathbb{P}\left(\cap_{j\in[p]}\{\xi_{ij} = k_j\}\right) \prod_{j\in[p]} z_j^{\xi_{ij}}, \tag{5.13}$$

and $\mathbf{h}(\mathbf{z}) = (h_1(\mathbf{z}), \ldots, h_p(\mathbf{z}))$.

Denote by $m_{ij} := \frac{\partial h_i}{\partial z_j}(\mathbf{1}) = \mathbb{E}[\xi_{ij}]$ and by $M = (m_{i,j})$ the mean matrix. We say that M is *finite* if $m_{i,j} < \infty$ for all i, j. We say that M is *irreducible* if for every pair i, j there exists $t \in \mathbb{N}$ such that $(M^t)_{i,j} > 0$. Let ρ be the largest eigenvalue of M. Let $\mathbf{u}$ and $\mathbf{v}$ be the right and left eigenvectors of M of eigenvalue ρ with the convention that $\mathbf{v}\mathbf{u}^\top = 1$ and $\mathbf{u}\mathbf{1}^\top = 1$.

A p-type branching process starting at $a \in [p]$ with offspring distribution Ξ is a

stochastic process

$$(\mathbf{X}^{(a)}(t) = (X_1^{(a)}(t), \ldots, X_p^{(a)}(t)))_{t\geq 0}, \tag{5.14}$$

defined as follows:

$$X_j^{(a)}(t) = \begin{cases} \mathbf{1}_{a=j} & (t=0) \\ \sum_{i\in[p]} \sum_{m=1}^{X_i^{(a)}(t-1)} \xi_{ij}(m; t-1) & (t \geq 1) \end{cases}$$

Note that

$$\mathbb{E}[\mathbf{X}^{(a)}(t)] = \mathbf{e}_a M^t \tag{5.15}$$

We call a process *irreducible* if its mean matrix is irreducible.

If $\mathbf{X}^{(a)}(t) \neq \mathbf{0}$ for all $t \in \mathbb{N}$, then the branching process is said to *survive*; otherwise, it is said to *become extinct*. Let $s^{(a)} = \mathbb{P}\left(\cap_{t\geq 0}\{\mathbf{X}^{(a)}(t) \neq \mathbf{0}\}\right)$ and $q^{(a)} = 1 - s^{(a)}$. Write $\mathbf{s} = (s^{(1)}, \ldots, s^{(p)})$ and $\mathbf{q} = (q^{(1)}, \ldots, q^{(p)}) = \mathbf{1} - \mathbf{s}$.

We will use the following results about p-type branching processes:

Lemma 5.2.1 (Theorem VIII.3.2 in [3]). *Suppose M is finite, irreducible and non-singular. Then,*

i) if $\rho \leq 1$, $\mathbf{q} = \mathbf{1}$;

ii) if $\rho > 1$, $\mathbf{q} < \mathbf{1}$ and $\mathbf{q}$ is the only solution of $\mathbf{h}(\mathbf{z}) = \mathbf{z}$ with $\mathbf{q} < \mathbf{1}$.

Lemma 5.2.2 (Theorem 2.1 in [38]). *Suppose that M is finite, strictly positive and nonsingular, and $\rho > 1$. Then there exists a sequence γ_t with*

$$\lim_{t\to\infty} (\gamma_t)^{1/t} = \rho^{-1}, \tag{5.16}$$

and non-negative random variables $W^{(a)}$ for $a \in [p]$ such that

$$\lim_{t\to\infty} \gamma_t \mathbf{X}^{(a)}(t) = W^{(a)}\mathbf{v}, \tag{5.17}$$

in probability. Moreover, $W^{(a)}$ is absolutely continuous in $(0,\infty)$ and $\mathbb{P}\left(W^{(a)}=0\right)=q^{(a)}$.

Remark 5.2.3. Note that the normalising sequence γ_t is independent of the starting type.

5.2.1 Supercritical processes conditional on extinction

In this section we study supercritical multitype branching processes conditional on extinction. Supercritical branching processes conditional on extinction are still branching processes [39] and we now describe the relevant parameters.

Let Ext_a denote the extinction event with an initial particle of type a; i.e., $\lim_{t\to\infty}\mathbf{X}^{(a)}(t)=\mathbf{0}$. We denote by Ext the extinction event, regardless of the initial type. Let $\hat{\mathbb{P}}(\cdot):=\mathbb{P}(\cdot\mid\mathrm{Ext})$. In order not to condition on an event of probability zero, we will impose $\mathbf{q}>\mathbf{0}$ throughout the section.

The conditioned process with law $\hat{\mathbb{P}}(\cdot)$ is a subcritical irreducible branching process $\hat{\mathbf{X}}^{(a)}(t)$ with offspring $\hat{\Xi}$ that has generating function

$$\hat{\mathbf{h}}(\mathbf{z})=\mathbf{q}^{-1}\mathbf{h}(\mathbf{q}\mathbf{z}) \tag{5.18}$$

where we slightly abuse the notations by letting

$$\mathbf{q}^{-1}:=(1/q^{(1)},\dots 1/q^{(p)}),\qquad\text{and}\qquad \mathbf{q}\mathbf{z}:=(q^{(1)}z^{(1)},\dots,q^{(p)}z^{(p)}). \tag{5.19}$$

Let $\hat{M}$ be the mean matrix of the process, which can be explicitly computed as follows:

$$\hat{m}_{i,j}=\frac{q^{(j)}}{q^{(i)}}\frac{\partial h_i}{\partial z_j}(\mathbf{q}). \tag{5.20}$$

Let $\hat{\rho}<1$ be the largest eigenvalue of $\hat{M}$.

We first prove the following result on supercritical p-type processes that do neither grow quickly nor become extinct.

Theorem 5.2.4. *Fix $p \in \mathbb{N}$ and $a \in [p]$. Let $(\mathbf{X}^{(a)}(t))_{t\geq 0}$ be an irreducible p-type branching process with offspring distribution Ξ and mean matrix M with $\rho \in (1,\infty)$ and $\mathbf{q} > \mathbf{0}$. Let γ_t be the normalising sequence in Lemma 5.2.2 and define $t_\omega := \inf\{t \geq 0 : (\gamma_{t'})^{-1} \geq \omega$ for all $t' \geq t\}$. Then there exist constants $c, C > 0$ depending on Ξ such that*

$$c\hat{\rho}^t \leq \mathbb{P}\left(\cap_{i=1}^t [\mathbf{0} \neq \mathbf{X}^{(a)}(i) < \boldsymbol{\omega}]\right) \leq C\hat{\rho}^{t-t_\omega}, \qquad \textit{for all } t \geq 1, \omega \geq t. \tag{5.21}$$

Proof. First we shall prove the upper bound. It suffices bound from above the probability that $\mathbf{X}^{(a)}(t) < \boldsymbol{\omega}$ conditional on $\mathbf{X}^{(a)}(t) \neq \mathbf{0}$. Define $r_t^{(a)} := \mathbb{P}\left(\mathbf{X}^{(a)}(t) < \boldsymbol{\omega} \mid \mathbf{X}^{(a)}(t) \neq \mathbf{0}\right)$ and $\mathbf{r}_t = (r_t^{(a)})_{a\in[p]}$. Let $Y^{(a)}(t) := \sum_{j\in[p]} X_j^{(a)}(t)$ be the random variable which tracks the total size of the t-th generation. Now, by definition of t_ω and absolute continuity of $W^{(a)}$,

$$\begin{aligned}
\mathbb{P}\left(\mathbf{X}^{(a)}(t_\omega) \not< \boldsymbol{\omega} \mid \mathbf{X}^{(a)}(t_\omega) \neq \mathbf{0}\right) &\geq \mathbb{P}\left(Y^{(a)}(t_\omega) \geq p\omega \mid \mathbf{X}^{(a)}(t) \neq \mathbf{0}\right) && (5.22)\\
&\geq \mathbb{P}\left(Y^{(a)}(t_\omega) \geq p\omega\right)\\
&\geq \mathbb{P}\left(Y^{(a)}(t_\omega) \geq p(\gamma_{t_\omega})^{-1}\right)\\
&\geq \mathbb{P}\left(p(\gamma_{t_\omega})^{-1} \leq Y^{(a)}(t_\omega) \leq 2p(\gamma_{t_\omega})^{-1}\right) > c_0^{(a)}\,, && (5.23)
\end{aligned}$$

where $c_0^{(a)} > 0$. So $r_{t_\omega}^{(a)} < 1 - c_0^{(a)}$ and $\mathbf{r}_{t_\omega} < \mathbf{1}$.

We claim that $r_t^{(a)}$ decreases exponentially in t with base approximately $\hat{\rho}$, for $t \geq t_\omega$. Define $r_t^{(\max)} := \max_{a\in[p]} r_t^{(a)}$. Let $\mathbf{Z}^{(a)}(t) \leq \mathbf{X}^{(a)}(1)$ be the vector giving the number of children of the initial particle that have progeny in the t-th generation, and write $z^{(a)}(t) = \sum_{j\in[p]} Z_j^{(a)}(t)$. The following holds,

$$r_t^{(a)} \leq \sum_{j\in[p]} \mathbb{P}\left(\mathbf{Z}^{(a)}(t) = \mathbf{e}_j \mid z^{(a)}(t) > 0\right) r_{t-1}^{(j)} + \mathbb{P}\left(z^{(a)}(t) \geq 2 \mid z^{(a)}(t) > 0\right) (r_{t-1}^{(\max)})^2\,. \tag{5.24}$$

Indeed, either there is only one child of the initial particle with surviving progeny at the

t-th generation and the probability that a child with type j has surviving progeny is r_{t-1}^{j}, or there are at least two children, each having probability of surviving progeny at most $r_{t-1}^{(\max)}$ independently from each other.

Let $\mathbf{o}_t := (\mathbb{P}\left(z^{(a)}(t) \geq 2 \mid z^{(a)}(t) > 0\right))_{a\in[p]}$ be the probability vector of having at least two children with progeny at time t. Moreover, define the matrix $\tilde{M}(t) = (\tilde{m}_{j,a}(t))$ by letting $\tilde{m}_{j,an}(t) := \mathbb{P}\left(\mathbf{Z}^{(a)}(t) = \mathbf{e}_j \mid z^{(a)}(t) > 0\right)$. Then, we can write (5.24) in matrix form as

$$\mathbf{r}_t \leq \mathbf{r}_{t-1}\tilde{M}(t) + (r_{t-1}^{(\max)})^2\mathbf{o}_t \ . \tag{5.25}$$

As $\mathbf{r}_{t_\omega} < \mathbf{1}$, $\mathbf{r}_t$ decreases in t. Thus to bound the growth rate of the vector $\mathbf{r}_t$ it suffices to bound the entries of $\tilde{M}(t)$. Define $s_t^{(a)} := \mathbb{P}\left(z^{(a)}(t) > 0\right)$ to be the probability the process survives up to time t and note that $s_t^{(a)} = 1 - h_a(\mathbf{1} - \mathbf{s}_t)$. Recall that Ext_a denotes the extinction event with initial particle of type a. The probability of the survival at time t satisfies the following,

$$\begin{aligned} s^{(a)} \leq s_t^{(a)} &= \mathbb{P}\left(\mathrm{Ext}_a^c\right) + \mathbb{P}\left(z^{(a)}(t) > 0, \mathrm{Ext}_a\right) \\ &= s^{(a)} + (1 - s^{(a)})\mathbb{P}\left(z^{(a)}(t) > 0 \mid \mathrm{Ext}_a\right) \\ &\leq s^{(a)} + (1 - s^{(a)})\hat{\rho}^t \ , \end{aligned} \tag{5.26}$$

where the final inequality follows by Markov's inequality and the fact that $\hat{\rho}$ is the leading eigenvalue of $\hat{M}$ as defined in (5.20).

If we condition on $\mathbf{X}^{(a)}(1)$, the events "a given particle x on the first generation has progeny at time t" are mutually independent each happening with probability $s_{t-1}^{(j)}$, where j is the type of x. Therefore, $\mathbf{Z}^{(a)}(t)$ is the $\mathbf{s}_{t-1}$-thinned version of $\mathbf{X}^{(a)}(1)$ (that is; for each $j \in [p]$, we consider that each children of the initial particle counted by $X_j^{(a)}(1)$ is also counted by $Z_j^{(a)}(t)$ independently with probability $s_{t-1}^{(j)}$) and so,

$$\tilde{m}_{j,a}(t) = \mathbb{P}\left(\mathbf{Z}^{(a)}(t) = \mathbf{e}_j \mid z^{(a)}(t) > 0\right) = \frac{\mathbb{P}\left(\mathbf{Z}^{(a)}(t) = \mathbf{e}_j\right)}{\mathbb{P}\left(z^{(a)}(t) > 0\right)} = \frac{s_{t-1}^{(j)}}{s_t^{(a)}}\frac{\partial h_a}{\partial z_j}(\mathbf{1} - \mathbf{s}_{t-1}) \tag{5.27}$$

Using the multivariable Taylor expansion and (5.26), we have

$$\frac{\partial h_a}{\partial z_j}(\mathbf{1}-\mathbf{s}_{t-1}) = \frac{\partial h_a}{\partial z_j}(\mathbf{1}-\mathbf{s}) + O(\hat{\rho}^t).$$

Substituting it into (5.27) and using (5.26) again allows us to deduce that

$$\tilde{m}_{j,a}(t) = \frac{s^{(j)}}{s^{(a)}}\frac{\partial h_a}{\partial z_j}(\mathbf{1}-\mathbf{s}) + O(\hat{\rho}^t) \tag{5.28}$$

Recall the definition of $\hat{M}$ in (5.20), and observe that $\tilde{M}(t) = P\hat{M}P^{-1} + O(\hat{\rho}^t)J$, where P is a diagonal matrix with entries $p_{aa} = q^{(a)}/s^{(a)}$ and J is the all-ones matrix. Thus the asymptotic behaviour of the eigenvalues of $\tilde{M}(t)$ is the same as the ones of $\hat{M}$. In particular its largest eigenvalue is $\hat{\rho}(1+O(\hat{\rho}^t))$. So, by (5.25) the vector $\mathbf{r}_t$ has exponential growth rate at most $\hat{\rho}(1+O(\hat{\rho}^t))$ as t goes to infinity, so

$$\mathbf{r}_t \leq \mathbf{r}_{t_\omega}\hat{\rho}^{t-t_\omega}\prod_{i=t_\omega}^{t}(1+O(\hat{\rho}^i)) \leq C\hat{\rho}^{t-t_\omega}\mathbf{1}\ ,$$

concluding the proof of the upper bound.

Next, we give a proof of the claimed lower bound. For $i \in [t]$, let $\tilde{\mathbf{X}}^{(a)}(i) \leq \mathbf{X}^{(a)}(i)$ be the subprocess of the elements that have some surviving progeny. For the lower bound, consider the following events:

$$E_1^{(a)} = [\tilde{\mathbf{X}}^{(a)}(t)\mathbf{1}^\top = 1], \qquad E_2^{(a)} = \left[\cap_{i=1}^t[\mathbf{0} \neq \mathbf{X}^{(a)}(i) < \boldsymbol{\omega}]\right]. \tag{5.29}$$

In words, $E_1^{(a)}$ is the event that, starting with an individual of type a, at time t there is exactly one particle that has surviving progeny.

Instead of bounding from below the probability of $E_2^{(a)}$, we will give a lower bound for the probability of $E^{(a)} := E_1^{(a)} \cap E_2^{(a)}$. Write

$$\mathbb{P}\left(E^{(a)}\right) = \mathbb{P}\left(E_1^{(a)}\right)\mathbb{P}\left(E_2^{(a)} \mid E_1^{(a)}\right). \tag{5.30}$$

We start by computing $\mathbb{P}\left(E_1^{(a)}\right)$. Conditioning on that an element of $\mathbf{X}^{(a)}(i)$ belongs to $\tilde{\mathbf{X}}^{(a)}(i)$ is equivalent to conditioning on that the progeny of at least one of its children survives. So, conditional on $\mathbf{X}^{(a)}(i)$ surviving, $\tilde{\mathbf{X}}^{(a)}(i)$ is a branching process with offspring distribution $\tilde{\Xi} = (\tilde{\boldsymbol{\xi}}_a)$, the $\mathbf{s}$-*thinned* version of Ξ, conditioned on being non-zero. Similarly to before, let $\tilde{m}_{j,a} := \mathbb{P}\left(\boldsymbol{\xi}_a = \mathbf{e}_j \mid \mathrm{Ext}_a^c\right) = \mathbb{P}\left(\tilde{\boldsymbol{\xi}}_a = \mathbf{e}_j\right)$. So, as in (5.28),

$$\tilde{m}_{j,a} = \frac{s^{(j)}}{s^{(a)}} \frac{\partial h_a}{\partial z_j}(\mathbf{1} - \mathbf{s}) \ ,$$

and $\tilde{M} = (\tilde{m}_{a,j}) = P\hat{M}P^{-1}$ has largest eigenvector $\hat{\rho}$. We can conclude that

$$\mathbb{P}\left(E_1^{(a)}\right) = \mathbb{P}\left(\mathrm{Ext}_a^c, \cap_{i=1}^t [\tilde{\mathbf{X}}^{(a)}(i)\mathbf{1}^\top = 1]\right) = s^{(a)}\mathbf{e}_a \tilde{M}^t \mathbf{1}^\top \tag{5.31}$$

Note that M is irreducible and hence so is $\tilde{M}$. By the Perron-Frobenius theorem, the largest eigenvalue, $\hat{\rho}$ of $\tilde{M}$ is simple and has associated left eigenvector $\tilde{\mathbf{v}}_1$ that is positive and satisfies $\tilde{\mathbf{v}}_1\mathbf{1}^\top = 1$. Moreover, there exist $\alpha^{(a)} > 0$ and $\beta_j^{(a)}$ such that we may write

$$\mathbf{e}_a = \alpha^{(a)}\tilde{\mathbf{v}}_1 + \sum_{j=2}^{p} \beta_j^{(a)} \tilde{\mathbf{v}}_j \ ,$$

where the $\tilde{\mathbf{v}}_j$ extend $\tilde{\mathbf{v}}_1$ to an orthogonal basis. Thus if λ_2 is the second largest eigenvalue of $\tilde{M}$ (in absolute value), $|\lambda_2| < \hat{\rho} < 1$ and we have

$$\begin{aligned}
\mathbf{e}_a \tilde{M}^t \mathbf{1}^\top &= \alpha^{(a)}\hat{\rho}^t \tilde{\mathbf{v}}_1 \mathbf{1}^\top + \sum_{j=2}^{k} \beta_j^{(a)} \tilde{\mathbf{v}}_j \tilde{M}^t \mathbf{1}^\top \\
&= \alpha^{(a)}\hat{\rho}^t \pm |\lambda_2|^t \left| \sum_{j=2}^{k} \beta_j^{(a)} \tilde{\mathbf{v}}_j \mathbf{1}^\top \right| \\
&= (1 + O((|\lambda_2|/\hat{\rho})^t))\alpha^{(a)}\hat{\rho}^t \ .
\end{aligned} \tag{5.32}$$

We may deduce from (5.31) that $\mathbb{P}\left(E_1^{(a)}\right) \geq (1 + o(1))s^{(a)}\alpha^{(a)}\hat{\rho}^t$.

When we condition on $E_1^{(a)}$, the tree is a spine $u_0, u_1, \ldots, u_t$ of length t where we attach at independent p-type branching processes to each u_i. For $i \in [t]$, and if a and

j are the types of u_i and u_{i+1} respectively, the process attached to u_i is constructed as follows: (i) the root u_i has offspring distribution $\boldsymbol{\xi}_a - \mathbf{e}_j$ conditional on $\boldsymbol{\xi}_a \geq \mathbf{e}_j$, and (ii) any other particle has offspring distribution $\hat{\Xi}$. Note that the number of children of u_i is stochastically dominated by $1 + \boldsymbol{\xi}_a \mathbf{1}^\top$, so it has finite expectation and variance; we will use it in the computations below.

By Markov's inequality, from (5.32) we can deduce that for $i \in [t]$,

$$\mathbb{E}(\mathbf{X}^{(a)}(i)\mathbf{1}^\top \mid E_1^{(a)}) = 1+O(1)\sum_{x=0}^{i-1}\mathbb{E}(\tilde{\mathbf{X}}^{(a)}(x)\mathbf{1}^\top) = 1+O(1)\sum_{x=0}^{i-1}\mathbf{e}_a\tilde{M}^x\mathbf{1}^\top \leq 1+\sum_{x=0}^{i-1}O(\hat{\rho}^x) = O(1)\ ,$$

where the first term $O(1)$ accounts for the expected number of children of each element of the spine.

Let $\tilde{C}^{(a)}(t) = (c^{(a)}_{j_1 j_2}(t))$ where $c^{(a)}_{j_1 j_2} = \mathbb{E}[\tilde{X}^{(a)}_{j_1}(t)\tilde{X}^{(a)}_{j_2}(t)]$, denote the second moment matrix of $\tilde{\mathbf{X}}^{(a)}(t)$. In particular $\tilde{C}^{(a)}(0) = \mathbf{e}_a^\top \mathbf{e}_a$ and $\tilde{C}^{(a)}(1)$ is the second moment matrix of $\tilde{\boldsymbol{\xi}}_a$. We can write (see e.g. (4.3) in Harris [34])

$$\tilde{C}^{(a)}(i) = (\tilde{M}^\top)^i C^{(a)}(0)\tilde{M}^i + \sum_{x=1}^{i}(\tilde{M}^\top)^{i-x}\left(\sum_{j\in[p]}\mathrm{Var}(\tilde{\boldsymbol{\xi}}_j)\mathbb{E}[\tilde{X}^{(a)}_j(x-1)]\right)\tilde{M}^{i-x}$$

We will estimate $\mathbf{1}\tilde{C}^{(a)}(i)\mathbf{1}^\top$. The first part can be simplified using (5.32)

$$\mathbf{1}(\tilde{M}^\top)^i C^{(a)}(0)\tilde{M}^i\mathbf{1}^T = (\mathbf{e}_a\tilde{M}^i\mathbf{1}^T)^T(\mathbf{e}_a\tilde{M}^i\mathbf{1}^T) = O(\hat{\rho}^{2i})$$

and writing $\mathrm{Var}_{\max}\tilde{\Xi} = \max_{j,j_1,j_2}(\mathrm{Var}(\tilde{\boldsymbol{\xi}}_j))_{j_1,j_2}$ each term of the second part is

$$\begin{aligned}\mathbf{1}(\tilde{M}^\top)^{i-x}\left(\sum_{j\in[p]}\mathrm{Var}(\tilde{\boldsymbol{\xi}}_j)\mathbb{E}[\tilde{X}^{(a)}_j(x-1)]\right)\tilde{M}^{i-x}\mathbf{1}^\top &= O(\hat{\rho}^x)(\tilde{M}^{i-x}\mathbf{1}^\top)^\top\left(\sum_{j\in[p]}\mathrm{Var}(\tilde{\boldsymbol{\xi}}_j)\right)(\tilde{M}^{i-x}\mathbf{1}^\top)\\ &= O(\hat{\rho}^{2i-x}\mathrm{Var}_{\max}\tilde{\Xi})\end{aligned}$$

Thus, we can bound the variance

$$\mathrm{Var}(\tilde{\mathbf{X}}^{(a)}(i)\mathbf{1}^\top) \leq \mathbf{1}\tilde{C}^{(a)}(i)\mathbf{1}^\top = O(\hat{\rho}^i \mathrm{Var}_{\max}\tilde{\Xi})$$

and compute the variance of the conditional process

$$\mathrm{Var}(\mathbf{X}^{(a)}(i)\mathbf{1}^\top \mid E_1^{(a)}) = O(1)\sum_{x=0}^{i-1}\mathrm{Var}(\tilde{\mathbf{X}}^{(a)}(x)\mathbf{1}^\top) = O(\mathrm{Var}_{\max}\tilde{\Xi}) = O(1)\ ,$$

The first term $O(1)$ comes from the offspring variance of the root in the spine, and the last equality follows from the fact that the tensor $\mathrm{Var}(\tilde{\Xi})$ can be computed as a function of $\mathbf{h}$ and its derivatives evaluated at $\mathbf{q} < \mathbf{1}$ and $\mathbf{h}$ has radius of convergence at least 1, so $\mathrm{Var}_{\max}\tilde{\Xi} < \infty$.

By Chebyshev's inequality and the choice of $\omega \geq t$,

$$\mathbb{P}\left((E_2^{(a)})^c \mid E_1^{(a)}\right) \leq \sum_{i=1}^{t}\mathbb{P}\left(\mathbf{X}^{(a)}(i)\mathbf{1}^\top \geq \omega \mid E_1^{(a)}\right) \leq \sum_{i=1}^{t}\frac{\mathrm{Var}(\mathbf{X}^{(a)}(i)\mathbf{1}^\top \mid E_1^{(a)})}{\omega^2} = O\big(t/\omega^2\big) = o(1). \tag{5.33}$$

Thus, we can conclude that

$$\mathbb{P}\left(E^{(a)}\right) \geq (1+o(1))s^{(a)}\alpha^{(a)}\hat{\rho}^t \geq c\hat{\rho}^t\ , \tag{5.34}$$

for some $c > 0$, which concludes the proof of the theorem. □

We will use the previous theorem to prove the following.

Lemma 5.2.5. *For $a \in [p]$, let $(\mathbf{X}^{(a)}(t))_{t\geq 0}$ be an irreducible p-type branching process with offspring distribution ξ and mean matrix M with $\rho \in (1,\infty)$. Let*

$$T_\omega^{(a)} := \inf\{t : \mathbf{X}^{(a)}(t) \not< \boldsymbol{\omega}\}. \tag{5.35}$$

Then for all $\varepsilon > 0$ and as $\omega \to \infty$, and letting $t_1 = (1+\varepsilon)\log_\rho \omega$

$$\mathbb{P}\left(T_\omega^{(a)} \le t_1\right) \to 1 - q^{(a)}, \tag{5.36}$$

and

$$\mathbb{P}\left(T_\omega^{(a)} \in (t_1, \infty)\right) \to 0. \tag{5.37}$$

Proof. We first prove (5.37). By Theorem 5.2.4, there exist constants $C > 0$ and $\hat{\rho} \in (0,1)$ such that for all $\varepsilon > 0$,

$$\mathbb{P}\left(T_\omega^{(a)} \in (t_1, \infty)\right) \le \mathbb{P}\left(\cap_{i=1}^{t_1}[\mathbf{0} \ne \mathbf{X}^{(a)}(i) < \boldsymbol{\omega}]\right) \le C\hat{\rho}^{(1+\varepsilon)\log_\rho \omega - (1+o(1))\log_\rho \omega - 1} = o(1), \tag{5.38}$$

where we used that $t_\omega = (1+o(1))\frac{\log \gamma_{t_\omega}^{-1}}{\log \rho} = (1+o(1))\log_\rho \omega$, by (5.16) and the definition of t_ω. Note that Theorem 5.2.4 can only be applied if $\mathbf{q} > \mathbf{0}$, however this is only necessary for the lower bound, not the upper one.

To prove (5.36) it suffices to show that $\mathbb{P}\left(T_\omega^{(a)} > t_1\right) \to q^{(a)}$, and by (5.38), we have

$$\begin{aligned}
\mathbb{P}\left(T_\omega^{(a)} > t_1\right) &= \mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) + \mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{0} \ne \mathbf{X}^{(i)}(t_1) < \boldsymbol{\omega}]\right) \\
&= \mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) + \mathbb{P}\left(\cap_{i=1}^{t_1}[\mathbf{0} \ne \mathbf{X}^{(a)}(i) < \boldsymbol{\omega}]\right) \\
&= \mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) + o(1).
\end{aligned} \tag{5.39}$$

Let $Y^{(a)}(t) = \sum_{i=0}^{t} \mathbf{X}^{(a)}(i)\mathbf{1}^\top$ be the total progeny up to time t. Recall that Ext_a is the event that $\mathbf{X}^{(a)}$ becomes extinct eventually, i.e. $\lim_{t\to\infty} Y^{(a)}(t) < \infty$. If $q^{(a)} = \mathbb{P}(\mathrm{Ext}_a) = 0$, then (5.39) is $o(1)$ and we are done. So let us assume that $q^{(i)} > 0$. Then

$$\begin{aligned}
\mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) &\le \mathbb{P}\left(\left[Y^{(a)}(t_1) \le p(1+t_1)\omega\right] \cap [\mathbf{X}^{(i)}(t_1) = \mathbf{0}]\right) \\
&\le \mathbb{P}\left(Y^{(a)}(t_1) \le p(1+t_1)\omega \;\middle|\; \mathrm{Ext}_a\right)\mathbb{P}(\mathrm{Ext}_a) \\
&\to \mathbb{P}(\mathrm{Ext}_a) = q^{(a)},
\end{aligned} \tag{5.40}$$

since the supercritical branching process conditioned on becoming extinct is a subcritical process and so it has finite total progeny a.s. [76] and $\omega \to \infty$.

To lower bound (5.39), note that $Y^{(a)}(t) < \omega$ implies $T_\omega^{(a)} > t$. Thus,

$$\begin{aligned}\mathbb{P}\left([T_\omega^{(a)} > t_1] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) &\geq \mathbb{P}\left([Y^{(a)}(t_1) < \omega] \cap [\mathbf{X}^{(a)}(t_1) = \mathbf{0}]\right) \\ &= \mathbb{P}\left(Y^{(a)}(t_1) < \omega\right) - \mathbb{P}\left([Y^{(a)}(t_1) < \omega] \cap [\mathbf{X}^{(a)}(t_1) \neq \mathbf{0}]\right).\end{aligned} \tag{5.41}$$

Using again the finite progeny argument, we have that

$$\mathbb{P}\left(Y^{(a)}(t_1) < \omega\right) \geq \mathbb{P}\left(Y^{(a)}(t_1) < \omega \mid \mathrm{Ext}_a\right)\mathbb{P}\left(\mathrm{Ext}_a\right) \to \mathbb{P}\left(\mathrm{Ext}_a\right) = q^{(a)}. \tag{5.42}$$

We will use Lemma 5.2.2 to conclude. By (5.16), we have that $t_1 \geq (1+\epsilon/2)\frac{\log \gamma_{t_1}^{-1}}{\log \rho}$, so $(\gamma_{t_1})^{-1} \geq \rho^{(1-\epsilon/2)t_1} \geq \omega^{1+\epsilon/2}$. Moreover, the chosen left eigenvector $\mathbf{v}$ is positive. Recall that $t_1 = t_1(\omega)$. For all $\delta > 0$ and by (5.17)

$$\begin{aligned}\lim_{\omega\to\infty} \mathbb{P}\left(Y^{(a)}(t_1) < \omega \mid \mathbf{X}^{(a)}(t_1) \neq \mathbf{0}\right) &\leq \lim_{\omega\to\infty} \mathbb{P}\left(\mathbf{X}^{(a)}(t_1) < \boldsymbol{\omega} \mid \mathbf{X}^{(a)}(t_1) \neq \mathbf{0}\right) \\ &= \lim_{\omega\to\infty} \frac{\mathbb{P}\left(\mathbf{0} \neq \mathbf{X}^{(a)}(t_1) < \boldsymbol{\omega}\right)}{\mathbb{P}\left(\mathbf{X}^{(a)}(t_1) \neq \mathbf{0}\right)} \\ &\leq \frac{\mathbb{P}\left(\mathbf{0} \neq W^{(a)}\mathbf{v} < \delta\mathbf{1}\right)}{1 - q^{(a)}}.\end{aligned}$$

Since $\delta > 0$ is arbitrary and $W^{(a)}$ is absolutely continuous in $(0, \infty)$, we have

$$\mathbb{P}\left(Y^{(a)}(t_1) < \omega \mid \mathbf{X}^{(a)}(t_1) \neq \mathbf{0}\right) \to 0. \tag{5.43}$$

Putting (5.43) and (5.42) into (5.41), and then putting it together with (5.40) into (5.39), gives the desired lower bound. □

The previous result can be generalized to multiple iid branching processes.

Corollary 5.2.6. *Let* $\mathbf{d} = (d_1, \ldots, d_p) \in \mathbb{Z}_+^p$ *and* $\{\mathbf{X}^{(a)}(t,1), \ldots, \mathbf{X}^{(a)}(t,d_a)\}_{a\in[p]}$ *be a collection of mutually independent irreducible p-type branching processes with offspring*

distribution Ξ *and mean matrix* M *with* $\rho \in (1, \infty)$.

Let

$$T_\omega^{(\mathbf{d})} := \inf\Big\{t : \sum_{a\in[p]} \sum_{d\in[d_a]} \mathbf{X}^{(a)}(t,d) \not< \boldsymbol{\omega}\Big\}. \tag{5.44}$$

Then for all $\varepsilon > 0$ *and* $\omega \to \infty$, *and letting* $t_1 = (1+\varepsilon)\log_\rho \omega$

$$\mathbb{P}\left(T_\omega^{(\mathbf{d})} \le t_1\right) \to 1 - \prod_{a\in[p]} (q^{(a)})^{d_a}. \tag{5.45}$$

and

$$\mathbb{P}\left(T_\omega^{(\mathbf{d})} \in (t_1, \infty)\right) \to 0. \tag{5.46}$$

Proof. For $a \in [p]$ and $d \in [d_a]$, let $T_\omega^{(a)}(d)$ be the stopping time defined in (5.35) for $\mathbf{X}^{(a)}(t,d)$. By Lemma 5.2.5

$$\begin{aligned}
\mathbb{P}\left(T_\omega^{(\mathbf{d})} > t_1\right) &\le \mathbb{P}\left(\cap_{a\in[p]} \cap_{d\in[d_a]} \{T_\omega^{(a)}(d) > t_1\}\right) \\
&= \prod_{a\in[p]} \prod_{d\in[d_a]} \mathbb{P}\left(T_\omega^{(a)}(d) > t_1\right) \\
&\to \prod_{a\in[p]} (q^{(a)})^{d_a},
\end{aligned}$$

and, writing $D = \sum_{a\in[p]} d_a$, $\omega' = \frac{\omega}{D}$, so $\log_\rho \omega' = \log_\rho \omega - O(1)$, we have

$$\begin{aligned}
\mathbb{P}\left(T_\omega^{(\mathbf{d})} > t_1\right) &\ge \mathbb{P}\left(\cap_{a\in[p]} \cap_{d\in[d_a]} \{T_{\omega'}^{(a)}(d) > t_1\}\right) \\
&= \prod_{a\in[p]} \prod_{d\in[d_a]} \mathbb{P}\left(T_{\omega'}^{(a)}(d) > t_1\right) \\
&\to \prod_{a\in[p]} (q^{(a)})^{d_a}.
\end{aligned}$$

This proves the first part of the lemma, the second part can be proven analogously using (5.37). □

5.3 Exploring the weak component

We will use a breath-first search (BFS) digraph exploration process on the weak components of $\mathbb{DCM}$ starting at a vertex $v \in [n]$. This is equivalent to the usual BFS process on the graph obtained by removing the directions of the edges in $\mathbb{DCM}$.

For $\mathcal{I} \subseteq [n]$, let $\mathcal{E}^{\pm}(\mathcal{I})$ be the set of heads/tails incident to the nodes in $\mathcal{I}$. Let $\mathcal{E}^{\pm} := \mathcal{E}^{\pm}([n])$. For $\mathcal{E} \subseteq \mathcal{E}^{\pm}$, let $\mathcal{V}(\mathcal{E})$ be the set of nodes incident to $\mathcal{E}$; we use $\mathcal{V}(e) = \mathcal{V}(\{e\})$. Let H be a partial pairing of half edges in $\mathcal{E}^{\pm}$. Let $\mathcal{P}^{\pm}(H) \subseteq \mathcal{E}^{\pm}$ be the set of heads/tails which are paired in H. Let $\mathcal{V}(H) = \mathcal{V}(\mathcal{P}^{\pm}(H))$. Let $\mathcal{F}^{\pm}(H) := \mathcal{E}^{\pm}(\mathcal{V}(H)) \setminus \mathcal{P}^{\pm}(H)$ be the unpaired heads/tails which are incident to $\mathcal{V}(H)$. Let E_H denote the event that H is a subgraph of $\mathbb{DCM}$. We will explore the graph conditioning on E_H.

We start from an arbitrary vertex $v \in [n] \setminus \mathcal{V}(H)$. In this process, we create random pairings of half-edges one by one and keep each half-edge in exactly one of the four states — *active, paired*, *fatal* or *undiscovered.* Let $\mathcal{A}_i^{\pm}$, $\mathcal{P}_i^{\pm}$, $\mathcal{F}_i^{\pm}$ and $\mathcal{U}_i^{\pm}$ denote the set of heads/tails in the four states respectively after the i-th pairing of half-edges, and $\mathcal{A}_i = \mathcal{A}_i^{-} \cup \mathcal{A}_i^{+}$. Initially, let

$$\mathcal{A}_0^{\pm} = \mathcal{E}^{\pm}(\{v\}),\ \mathcal{P}_0^{\pm} = \mathcal{P}^{\pm}(H),\ \mathcal{F}_0^{\pm} = \mathcal{F}^{\pm}(H),\ \mathcal{U}_0^{\pm} = \mathcal{E}^{\pm} \setminus (\mathcal{A}_0^{\pm} \cup \mathcal{P}_0^{\pm} \cup \mathcal{F}_0^{\pm}). \tag{5.47}$$

Then set $i = 1$ and proceed as follows:

(i) Let e_i^* be one of the half-edges that became active earliest in $\mathcal{A}_{i-1}$ with $* \in \{-, +\}$ and let $\sharp \in \{-, +\} \setminus \{*\}$.

(ii) Pair e_i^* with a half-edge $f_i^{\sharp}$ chosen uniformly at random from $\mathcal{E}^{\sharp} \setminus \mathcal{P}_{i-1}^{\sharp}$. Let $\mathcal{P}_i^* = \mathcal{P}_{i-1}^* \cup \{e_i^*\}$ and $\mathcal{P}_i^{\sharp} = \mathcal{P}_{i-1}^{\sharp} \cup \{f_i^{\sharp}\}$.

(iii) If $f_i^{\sharp} \in \mathcal{F}_{i-1}^{\sharp}$, then terminate; otherwise if $f_i^{\sharp} \in \mathcal{A}_{i-1}^{\sharp}$, then $\mathcal{A}_i^* = \mathcal{A}_{i-1}^* \setminus \{e_i^*\}$ and $\mathcal{A}_i^{\sharp} = \mathcal{A}_{i-1}^{\sharp} \setminus \{f_i^{\sharp}\}$; and if $f_i^{\sharp} \in \mathcal{U}_{i-1}^{\sharp}$, then $\mathcal{A}_i^* = (\mathcal{A}_{i-1}^* \cup \mathcal{E}^*(v_i)) \setminus \{e_i^*\}$ and $\mathcal{A}_i^{\sharp} = (\mathcal{A}_{i-1}^{\sharp} \cup \mathcal{E}^{\sharp}(v_i)) \setminus \{f_i^{\sharp}\}$, where $v_i = \mathcal{V}(f_i^{\sharp})$.

(iv) If $\mathcal{A}_i = \emptyset$ terminate; otherwise, $\mathcal{F}_i^{\pm} = \mathcal{F}_{i-1}^{\pm}$, $\mathcal{U}_i^{\pm} = \mathcal{E}^{\pm} \setminus (\mathcal{A}_i^{\pm} \cup \mathcal{P}_i^{\pm} \cup \mathcal{F}_i^{\pm})$, $i = i+1$ and go to (i).

Let $F_v(0)$ be a rooted forest composed of $|\mathcal{E}^-(v)| + |\mathcal{E}^+(v)|$ isolated nodes (called roots) corresponding to the half-edges incident to v. Given $F_v(i-1)$, $F_v(i)$ is constructed as follows: if $f_i^\sharp \in \mathcal{U}_{i-1}^\sharp$, then construct $F_v(i)$ from $F_v(i-1)$ by adding $\left|\mathcal{E}^{\pm}(v_i) \setminus \{f_i^\sharp\}\right|$ child nodes to the node representing e_i^*, each of which represents a half-edge incident to v_i different from $f_i^\sharp$; otherwise, let $F_v(i) = F_v(i-1)$. For each $e \in \mathcal{E}^{\pm}(v)$, we denote by $F_e(i)$ the tree corresponding to e at time i. We assign two labels to each non-root node. First, we assign the label *head* if the node corresponded to a head in $\mathcal{E}^-$, and *tail* otherwise. Second, as all nodes correspond to the half-edges in $(\mathcal{P}_i^{\pm} \setminus \mathcal{P}_0^{\pm}) \cup \mathcal{A}_i^{\pm}$, we assign the label *active* or *paired* accordingly.

If i_t is the last step where a half-edge at *undirected* distance t from v is paired, then $F_v(i_t)$ satisfies: (i) the height is t, as a rooted forest; (ii) the set of actives nodes is the t-th level. We call a rooted forest F *incomplete* if it satisfies (i)-(ii). We let $p(F)$ be the number of *paired* nodes in F.

5.3.1 Size biased distributions

We recall some notation in [11]. The *in-* and *out-size biased* distributions of D_n and D are defined by

$$\mathbb{P}\left((D_n)_{\text{in}} = (k-1,\ell)\right) = \frac{kn_{k,\ell}}{m_n}, \qquad \mathbb{P}\left((D_n)_{\text{out}} = (k,\ell-1)\right) = \frac{\ell n_{k,\ell}}{m_n}, \tag{5.48}$$

$$\mathbb{P}\left(D_{\text{in}} = (k-1,\ell)\right) = \frac{k\lambda_{k,\ell}}{\lambda}, \qquad \mathbb{P}\left(D_{\text{out}} = (k,\ell-1)\right) = \frac{\ell\lambda_{k,\ell}}{\lambda}. \tag{5.49}$$

Then, by (i) of Condition 5.1.1, $(D_n)_{\text{in}} \to D_{\text{in}}$ and $(D_n)_{\text{out}} \to D_{\text{out}}$ in distribution. Consider the sequence of random matrices $(\Xi_n)_{n\geq 0}$, defined by

$$\Xi_n = \begin{pmatrix} (D_n)^-_{\text{out}} & (D_n)^+_{\text{out}} \\ (D_n)^-_{\text{in}} & (D_n)^+_{\text{in}} \end{pmatrix} \tag{5.50}$$

with mean matrix

$$M_n = \frac{1}{\lambda} \begin{pmatrix} \mathbb{E}[D_n^- D_n^+] & \mathbb{E}[(D_n^+)^2] \\ \mathbb{E}[(D_n^-)^2] & \mathbb{E}[D_n^- D_n^+] \end{pmatrix} \tag{5.51}$$

having largest eigenvalue ρ_n and largest left eigenvalue $\mathbf{v}_n$, with the usual convention. Let $\mathbf{q}_n = (q_n^-, q_n^+)$ be the extinction probability vector for the 2-type process with offspring Ξ_n.

By Condition 5.1.1, we have

(i) Ξ_n converges in distribution to Ξ as defined in (5.5);

(ii) M_n converges to M as defined in (5.6) and both are finite and irreducible;

(iii) ρ_n converges to ρ, the largest eigenvalue of M;

(iv) $\mathbf{q}_n = (q_n^-, q_n^+)$ converges to $\mathbf{q} = (q^-, q^+)$, the extinction probability vector for the 2-type process with offspring Ξ.

Moreover, both M_n and M are finite and irreducible.

5.3.2 Coupling with branching processes

We will define two new sequences $\Xi_n^\uparrow$ and $\Xi_n^\downarrow$ that, asymptotically, will stochastically dominate and be stochastically dominated by Ξ_n. This will allow to couple the digraph exploration process with the 2-type branching processes defined in Section 5.2.

Write $X = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{pmatrix}$ and $Z = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ the zero matrix. Let $x_{\max} = \max_{1 \le i,j \le 2}\{x_{i,j}\}$. For $\beta > 0$ sufficiently small, consider the distributions $\Xi_n^\downarrow = \Xi_n^\downarrow(\beta)$ and $\Xi_n^\uparrow = \Xi_n^\uparrow(\beta)$ de-

fined by

$$\mathbb{P}\left(\Xi_n^{\downarrow} = X\right) = \begin{cases} c^{\downarrow}\mathbb{P}\left(\Xi_n = X\right) & \text{if } \mathbb{P}\left(\Xi_n = X\right) \geq n^{-2\beta} \text{and } x_{\max} \leq n^{\beta} \\ 0 & \text{otherwise} \end{cases} \tag{5.52}$$

$$\mathbb{P}\left(\Xi_n^{\uparrow} = X\right) = \begin{cases} c^{\uparrow}\mathbb{P}\left(\Xi_n = X\right) & X \neq Z \\ c^{\uparrow}\mathbb{P}\left(\Xi_n = Z\right) + n^{-1/2+2\beta} & X = Z \end{cases} \tag{5.53}$$

where $c^{\downarrow}$ and $c^{\uparrow}$ are normalising constants.

Convergence of the seconds moment of D_n in Condition 5.1.1 implies that $c^{\downarrow}, c^{\uparrow} = 1 + O(n^{-\beta})$, $M_n^{\uparrow}, M_n^{\downarrow} = (1 + o(1))M_n$ having largest eigenvalues $\rho_n^{\uparrow}, \rho_n^{\downarrow} = (1 + o(1))\rho_n$ and that $q_n^{\pm,\uparrow}, q_n^{\pm,\downarrow} = (1 + o(1))q_n^{\pm}$. Thus, these new distributions mimic the asymptotic behaviour of Ξ_n in distribution and in mean.

Let $\mathrm{GW}_{\Xi}^{(d_1,d_2)} = (\mathrm{GW}_{\Xi}^{(1)}(1), \ldots, \mathrm{GW}_{\Xi}^{(1)}(d_1); \mathrm{GW}_{\Xi}^{(2)}(1), \ldots, \mathrm{GW}_{\Xi}^{(2)}(d_2))$; be $d_1 + d_2$ independent 2-type Galton-Watson trees with offspring distribution Ξ, the first d_1 ones starting with a particle of type 1 (which we associate to heads), and the last d_2, starting with a particle of type 2 (associated to tails). Let $F = (T^{(1)}(1), \ldots, T^{(1)}(d_1); T^{(2)}(1), \ldots, T^{(2)}(d_2))$ be an incomplete rooted forest with $d_1 + d_2$ components. We denote by $\mathrm{GW}_{\Xi}^{(d_1,d_2)} \cong F$ the event that for each $j \in \{1, 2\}$ and $i \in [d_j]$,

i) $T^{(j)}(i)$ is a root subtree of $\mathrm{GW}_{\Xi}^{(j)}(i)$;

ii) all paired nodes of $T^{(j)}(i)$ have the same degree in $\mathrm{GW}_{\Xi}^{(j)}(i)$; and

iii) elements of $\mathrm{GW}_{\Xi}^{(j)}(i)$ of type 1 and 2 correspond to nodes labelled as heads and tails in $T^{(j)}(i)$, respectively.

We will need the following coupling lemma, which is a 2-dimensional version of [11, Lemma 5.3]. We omit its proof as it follows from an analogous argument.

Lemma 5.3.1. *Let $\beta > 0$ be sufficiently small and let H be a partial pairing with $|\mathcal{V}(H)| \leq n^{1-6\beta}$. Let $v \in [n] \setminus \mathcal{V}(H)$ with $d_v^- = d_1$ and $d_v^+ = d_2$. For every incom-*

plete forest F with $p(F) \leq n^\beta$ paired nodes, we have

$$(1-o(1))\mathbb{P}\left(\mathrm{GW}^{(d_1,d_2)}_{\Xi^{\downarrow}_n(\beta)} \cong F\right) \leq \mathbb{P}\left(F_v(p(F)) = F \mid E_H\right) \leq (1+o(1))\mathbb{P}\left(\mathrm{GW}^{(d_1,d_2)}_{\Xi^{\uparrow}_n(\beta)} \cong F\right). \tag{5.54}$$

5.4 Expansion probability

Given $x, y \in [n]$ an *undirected walk* of length t is a sequence

$$(x = v_0, e_0), (f_1, v_1, e_1), \ldots, (f_{t-1}, v_{t-1}, e_{t-1}), (f_t, v_t = y)$$

such that $v_i \in [n]$ for $i \in \{0, \ldots, t\}$, $e_i, f_i \in \mathcal{E}^{\pm}(v_i)$ and e_{j-1} is paired with f_j for all $j \in [t]$.

We say that a half-edge f is at *undirected distance* t from $v \in [n]$ if the shortest undirected walk from v to $\mathcal{V}(f)$ has length t. If it is clear from the context, we use distance instead of undirected distance. Denote by $\mathcal{N}^{\pm}_t(v)$ (and $\mathcal{N}^{\pm}_{\leq t}(v)$) the set of head/tails at distance (at most) t from v.

Fix

$$\omega := \log^6 n, \qquad t_0 := \log_\rho \omega = O(\log\log n). \tag{5.55}$$

Let $t_\omega(v)$ be the *expansion time* of v defined as

$$t_\omega(v) := \inf\left\{t \geq 1 : \max\{\left|\mathcal{N}^-_t(v)\right|, \left|\mathcal{N}^+_t(v)\right|\} \geq \omega\right\}. \tag{5.56}$$

Given H a partial pairing of $\mathcal{E}^{\pm}$ and $\mathcal{I} \subseteq [n]$, we consider the following event:

$$A_1(v, \varepsilon) := [t_\omega(v) \leq (1+\varepsilon)t_0]. \tag{5.57}$$

The first lemma in this section shows that the probability this event happens is close to the survival probability of a branching process.

Lemma 5.4.1. *Assume that $\rho > 1$. Fix $k, \ell \in N$, $\varepsilon \in (0, 1/2)$ and β sufficiently small. Then uniformly for all choices of partial pairing H and $v \in [n]$ that satisfy $|\mathcal{V}(H)| \leq n^\beta$ and $(d_v^-, d_v^+) = (k, \ell)$, as $n \to \infty$,*

$$\mathbb{P}\left(A_1(v, \varepsilon) \mid E_H\right) = (1 + o(1))(1 - (q^-)^k (q^+)^\ell). \tag{5.58}$$

Proof. Let $t_1 = \lfloor (1 + \varepsilon) t_0 \rfloor$. Define

$$A_2(v, H) := \left[\mathcal{N}^-_{\leq t_\omega(v) \wedge t_1}(v) \cap \mathcal{F}^-(H) = \emptyset\right] \cap \left[\mathcal{N}^+_{\leq t_\omega(v) \wedge t_1}(v) \cap \mathcal{F}^+(H) = \emptyset\right]$$

In words, this is the event that the in- or out- neighbourhood of v intersects H before the explosion time and before t_1.

Let $\mathfrak{F}_{k,\ell,t,\omega}$ be the class of incomplete rooted forests F with k trees having root labelled by head and ℓ trees having root labelled by tail, of height at most t and such that all levels have less than ω nodes with the same label (here we understand 'level' as the set all nodes at a given distance from any of the roots). For $F \in \mathfrak{F}_{k,\ell,t,\omega}$, we have $p(F) \leq \omega t_1 = O(\log^7 n)$. Let $\beta = \gamma/100$.

Let $(\mathbf{X}^{(1)}(t, 1), \ldots, \mathbf{X}^{(1)}(t, d_1); \mathbf{X}^{(2)}(t, 1), \ldots, \mathbf{X}^{(2)}(t, d_2))$ be independent 2-type branching processes with offspring distribution $\Xi_n^\downarrow(\beta)$ and extinction probability vector $(q_n^{(1),\downarrow}, q_n^{(2),\downarrow})$.

Recall that $q_n^{(1),\downarrow} \to q^- < 1$ and $q_n^{(2),\downarrow} \to q^+ < 1$ and the definition of $T_\omega^{(d_1,d_2)}$ given

in (5.44). By Corollary 5.2.6 and Lemma 5.3.1, the LHS of (5.58) is

$$\begin{aligned}
\mathbb{P}\left(A_1(v,\varepsilon)\cap A_2(v,H)\mid E_H\right) &= 1-\sum_{m=0}^{\lfloor \omega t\rfloor}\sum_{\substack{F\in\mathfrak{F}_{k,\ell,t,\omega}\\ p(F)=m}}\mathbb{P}\left(F_v(m)=F\mid E_H\right)\\
&\geq 1-(1+o(1))\sum_{m=0}^{\lfloor \omega t\rfloor}\sum_{\substack{F\in\mathfrak{F}_{k,\ell,t,\omega}\\ p(F)=m}}\mathbb{P}\left(\mathrm{GW}^{(k,\ell)}_{\Xi_n^{\downarrow}(\beta)}\cong F\right)\\
&= 1-(1+o(1))\mathbb{P}\left(T_\omega^{(k,\ell)}>t_1\right)\\
&= 1-(1+o(1))\left(1-\mathbb{P}\left(T_\omega^{(k,\ell)}\leq t_1\right)\right)\\
&= 1-(1+o(1))\left(1-(1-(q_n^{(1),\downarrow})^k(q_n^{(2),\downarrow})^\ell)\right)\\
&= (1+o(1))(1-(q^-)^k(q^+)^\ell).
\end{aligned}\tag{5.59}$$

where we used that $q^{\pm}<1$ in the last equality. The analogous lower bound follows from a similar argument, using $\Xi_n^{\uparrow}(\beta)$ instead.

It suffices to show that $A_2(v,H)$ is a likely event. Recall that Condition 5.1.1 implies that $\Delta=\max\{\Delta^-,\Delta^+\}=o(n^{1/2})$. At step i of the BFS exploration process, the probability that $f_i^\sharp\in\mathcal{F}_i^\sharp=\mathcal{F}^\sharp(H)$ is upper bounded by

$$p\coloneqq\frac{|\mathcal{F}_i^\sharp|}{m_n-|\mathcal{P}_i^\sharp|}\leq\frac{\Delta|\mathcal{V}(H)|}{m_n-i-|\mathcal{P}_0^\sharp|}=o(n^{-1/2+\beta})\ ,$$

as at most $K\coloneqq\omega t_1=O(\log^7 n)$ edges are paired when exposing $A_1(v,\varepsilon)$. Thus, the probability that at least one forbidden half-edge is paired is at most the probability that a binomial random variable with parameters K and p is at least 1. Thus,

$$\mathbb{P}\left((A_2(v,H))^c\right)\leq\mathbb{P}\left(\mathrm{Bin}(K,p)\geq 1\right)\leq Kp=o(1)\ .\tag{5.60}$$

Combining (5.59) and (5.60), we conclude

$$\begin{aligned}\mathbb{P}\left(A_1(v,\varepsilon) \mid E_H\right) &= \mathbb{P}\left(A_1(v,\varepsilon) \cap A_2(v,H) \mid E_H\right) + O\left(\mathbb{P}\left((A_2(v,H))^c \mid E_H\right)\right) \\ &= (1+o(1))(1-(q^-)^k(q^+)^\ell).\end{aligned}$$

where we use again that $q^\pm < 1$. □

The next result follows immediately by applying Lemma 5.4.1 twice, once on u with H the empty graph, and once on v with H being the neighbourhood of u generated up to time $t_\omega(u) \wedge (1+\varepsilon)t_0$, which is small by definition.

Corollary 5.4.2. *Assume that $\rho > 1$. Fix $\varepsilon \in (0,1/2)$ and distinct $u, v \in [n]$. As $n \to \infty$,*

$$\mathbb{P}\left(A_1(u,\varepsilon) \cap A_1(v,\varepsilon)\right) = (1+o(1))(1-(q^-)^{d_u^-}(q^+)^{d_u^+})(1-(q^-)^{d_v^-}(q^+)^{d_v^+}). \tag{5.61}$$

The following lemma shows that expansions are unlikely to happen very late.

Lemma 5.4.3. *Assume that $\rho > 1$. Fix $k, \ell \in N$, $\varepsilon \in (0,1/2)$. Then uniformly for all choices of $v \in [n]$ with $(d_v^-, d_v^+) = (k,\ell)$ and as $n \to \infty$,*

$$\mathbb{P}\left(\cap_{i=1}^t [0 < |\mathcal{N}_i^*(v)| < \omega], * \in \{-,+\}\right) \le \hat{\rho}^{t-O(\log\log n)} \qquad \textit{for any } t \ge 1. \tag{5.62}$$

In particular,

$$\mathbb{P}\left(t_\omega(v) \in ((1+\varepsilon)t_0, \infty)\right) = o(1). \tag{5.63}$$

Proof. Let $\omega' = \omega/(k+\ell)$. We have

$$\mathbb{P}\left(\cap_{i=1}^t [0 < |\mathcal{N}_i^*(v)| < \omega], * \in \{-,+\}\right) \le \sum_{e \in \mathcal{E}^\pm(v)} \mathbb{P}\left(\cap_{i=1}^t [0 < |\mathcal{N}_i^*(e)| < \omega'], * \in \{-,+\}\right) \tag{5.64}$$

where $\mathcal{N}_t^\pm(e)$ is defined as the set of head/tails at distance t from e.

We use the same ideas as in the proof of Lemma 5.4.1, without the need to condition on E_H and only for one tree. Let $\mathfrak{T}^{\pm}_{t,\omega}$ be the class of incomplete rooted trees T having root labelled by head/tail, of height t and such that every level has less than ω nodes of each label. For any head $e \in \mathcal{E}^-(v)$ (similarly for any tail in $\mathcal{E}^+(v)$), using Theorem 5.2.4 and Lemma 5.3.1 we have that

$$\begin{aligned}
\mathbb{P}\left(\cap_{i=1}^{t}[0 < |\mathcal{N}_i^*(e)| < \omega'], * \in \{-,+\}\right) &= \sum_{m=t-1}^{\lfloor \omega' t \rfloor} \sum_{\substack{T \in \mathfrak{T}^-_{t,\omega'} \\ p(F)=m}} \mathbb{P}\left(F_e(m) = T\right) \\
&\leq (1+o(1)) \sum_{m=t-1}^{\lfloor \omega' t \rfloor} \sum_{\substack{T \in \mathfrak{T}^-_{t,\omega'} \\ p(F)=m}} \mathbb{P}\left(\mathrm{GW}^{(1)}_{\Xi^{\uparrow}_n(\beta)} \cong T\right) \\
&= (1+o(1))\mathbb{P}\left(\cap_{i=1}^{t}[\mathbf{0} \neq \mathbf{X}^{(1)}(i) < \boldsymbol{\omega}]\right) \\
&\leq (1+o(1))C\hat{\rho}^{t-t_{\omega'}} \\
&= \hat{\rho}^{t-O(\log\log n)}.
\end{aligned} \tag{5.65}$$

where we used that $t_{\omega'} = (1+o(1))\log_\rho \omega' = (1+o(1))\log_\rho \omega = O(\log\log n)$. The proof follows by plugging the last bound into (5.64), and noting that $k, \ell = O(1)$. For (5.63), we set $t = (1+\varepsilon)t_0$ in (5.65) and conclude similarly. Alternatively, we can conclude from (5.46). □

Next, we estimate the number of vertices that have finite expansion time.

Proposition 5.4.4. *Assume that $\rho > 1$. Let*

$$\mathcal{L}_V := \{v \in [n] : t_\omega(v) < \infty\} \quad \textit{and} \quad \mathcal{L}_E := \left\{e \in \mathcal{E}^+ : v(e) \in \mathcal{L}_V\right\}. \tag{5.66}$$

Then

$$\frac{\mathbb{E}[|\mathcal{L}_V|]}{n} \to \eta, \quad \frac{\mathbb{E}[|\mathcal{L}_V|^2]}{n^2} \to \eta^2, \quad \frac{\mathbb{E}[|\mathcal{L}_E|]}{n} \to \zeta, \quad \textit{and} \quad \frac{\mathbb{E}[|\mathcal{L}_E|^2]}{n^2} \to \zeta^2 \tag{5.67}$$

where η and ζ are defined as in (5.9)*. Thus, $|\mathcal{L}_V|/n \to \eta$ and $|\mathcal{L}_E|/n \to \zeta$ in probability.*

Proof. As $\mathbf{q} < \mathbf{1}$ and $\sum_{k,\ell\geq 0} \lambda_{k,\ell} = 1$, we have $\eta \in (0,1]$; similarly $\zeta \in (0,\lambda]$. Fix $\varepsilon \in (0,1/2)$ and define

$$\mathcal{L}_V(\varepsilon) := \{v \in [n] : t_\omega(v) < (1+\varepsilon)t_0\} \quad \text{and} \quad \mathcal{L}_E(\varepsilon) := \left\{e \in \mathcal{E}^+ : v(e) \in \mathcal{L}_V(\varepsilon)\right\}. \tag{5.68}$$

which satisfy $\mathcal{L}_V(\varepsilon) \subseteq \mathcal{L}_V$ and $\mathcal{L}_E(\varepsilon) \subseteq \mathcal{L}_E$.

Given $v \in [n]$ with $(d_v^-, d_v^+) = (k,\ell)$, by Lemma 5.4.1 with H the empty graph

$$p_{k,\ell} := \mathbb{P}\left(v \in \mathcal{L}_V(\varepsilon)\right) = \mathbb{P}\left(A_1(v,\varepsilon)\right) = (1+o(1))(1-(q^-)^k(q^+)^\ell). \tag{5.69}$$

Since there are $n_{k,\ell}$ such nodes, by (i) of Condition 5.1.1 and using the dominated convergence theorem, we have

$$\frac{\mathbb{E}[|\mathcal{L}_V(\varepsilon)|]}{n} = \sum_{k,\ell\geq 0} \frac{n_{k,\ell}}{n} p_{k,\ell} = (1+o(1)) \sum_{k,\ell\geq 0} \frac{n_{k,\ell}}{n}(1-(q^-)^k(q^+)^\ell) \to \eta, \tag{5.70}$$

$$\frac{\mathbb{E}[|\mathcal{L}_E(\varepsilon)|]}{n} = \sum_{k,\ell\geq 0} \frac{kn_{k,\ell}}{n} p_{k,\ell} = (1+o(1)) \sum_{k,\ell\geq 0} \frac{kn_{k,\ell}}{n}(1-(q^-)^k(q^+)^\ell) \to \zeta. \tag{5.71}$$

On the other hand, Lemma 5.4.3 implies that $\mathbb{E}[|\mathcal{L}_V \setminus \mathcal{L}_V(\varepsilon)|] = o(n)$ and by Condition 5.1.1, the number of tails incident to $\mathcal{L}_V \setminus \mathcal{L}_V(\varepsilon)$ is $o(n)$, so $\mathbb{E}[|\mathcal{L}_E \setminus \mathcal{L}_E(\varepsilon)|] = o(n)$. This completes the proof of the expected values in (5.67).

For the second moment, choose distinct $u, v \in [n]$. By Corollary 5.4.2 we obtain

$$\begin{aligned}
\frac{\mathbb{E}[|\mathcal{L}_V(\varepsilon)|^2]}{n^2} &= \frac{1}{n^2}\big(\mathbb{E}[|\mathcal{L}_V(\varepsilon)|] + (1+o(1)) \sum_{u\neq v}(1-(q^-)^{d_u^-}(q^+)^{d_u^+})(1-(q^-)^{d_v^-}(q^+)^{d_v^+})\big) \\
&= o(1) + (1+o(1)) \sum_{k,\ell,k',\ell'\geq 0} \frac{n_{k,\ell}n_{k',\ell'}}{n^2}(1-(q^-)^k(q^+)^\ell)(1-(q^-)^{k'}(q^+)^{\ell'}) \\
&\to \eta^2.
\end{aligned} \tag{5.72}$$

The same argument shows that $\mathbb{E}[|\mathcal{L}_E(\varepsilon)|^2]/n^2 \to \zeta^2$. Also note that

$$\mathbb{E}[|\mathcal{L}_V|^2 - |\mathcal{L}_V(\varepsilon)|^2] \leq 2n\mathbb{E}[|\mathcal{L}_V \setminus \mathcal{L}_V(\varepsilon)|] = o(n^2).$$

and similarly $\mathbb{E}[|\mathcal{L}_E|^2 - |\mathcal{L}_E(\varepsilon)|^2] = o(n^2)$. This concludes the proof of the second moments in of (5.67). □

5.5 Connectivity of large sets of edges

In this section we show that any reasonably large pair of sets of edges will have a path between them. The main result of this section is the following.

Proposition 5.5.1. *Uniformly over all choices $\varepsilon, \gamma > 0$, partial pairings H and sets $\mathcal{X}, \mathcal{Y} \subseteq \mathcal{E}^{\pm} \setminus \mathcal{P}^{\pm}(H)$ such that $|\mathcal{V}(H)| < n^{1-\gamma}$ and $|\mathcal{X}|, |\mathcal{Y}| \geq \log^2 n$,*

$$\mathbb{P}\left(\mathrm{dist}(\mathcal{X}, \mathcal{Y}) > (1+\varepsilon)\log_\rho n \mid E_H\right) = o(n^{-2}). \tag{5.73}$$

It will suffice to find a path from $\mathcal{X}$ to $\mathcal{Y}$ constrained to only use vertices of bounded degree. To this end, for $\alpha > 0$ which we shall choose later, let $K \in \mathbb{N}$ be sufficiently large such that

$$\left(1 - \frac{\alpha}{4}\right)\mathbb{E}(D_{\mathrm{in}}^-) \leq \mathbb{E}(D_{\mathrm{in}}^- \mathbb{1}(D_{\mathrm{in}}^- < K,\ D_{\mathrm{in}}^+ < K))) \leq \mathbb{E}(D_{\mathrm{in}}^-) \tag{5.74}$$

with the same holding true with $D_{\mathrm{in}}^+, D_{\mathrm{out}}^-, D_{\mathrm{out}}^+$ in place of D_{in}^-. By Condition 5.1.1, it is straightforward to check that such a K exists.

We split $\mathcal{X}$ into $\mathcal{X} = \mathcal{X}^- \cup \mathcal{X}^+$ where $\mathcal{X}^- := \mathcal{X} \cap \mathcal{E}^-$, $\mathcal{X}^+ := \mathcal{X} \cap \mathcal{E}^+$. Moreover, let $\mathcal{L}^{\pm}$ be the set of heads and tails incident to vertices with at least K heads or K tails. Let $\mathcal{N}_0^+(\mathcal{X}) := \mathcal{X}^+$, $\mathcal{N}_0^-(\mathcal{X}) := \mathcal{X}^-$ and for $t \geq 1$ we define recursively,

$$\begin{aligned}\mathcal{N}_t^+(\mathcal{X}) = \{&e^+ \in \mathcal{E}^+ \setminus (\mathcal{L}^+ \cup \mathcal{N}_{<t}^+(\mathcal{X}) \cup \mathcal{F}^+(H)) \\ &\mid \exists f^* \in \mathcal{N}_{t-1}^*(\mathcal{X}^+), f^\sharp \in \mathcal{E}^\sharp(v(e^+)), f^\sharp \neq e^+, f^* f^\sharp \text{ is a pair}\}\end{aligned} \tag{5.75}$$

Similarly, we can define $\mathcal{N}_t^-(\mathcal{X})$.

Lemma 5.5.2. *Uniformly over all choices $\varepsilon, \gamma, \alpha > 0$, partial pairings H and $\mathcal{X} \subseteq \mathcal{E}^\pm \backslash \mathcal{P}^\pm(H)$ such that $|\mathcal{V}(H)| < n^{1-\gamma}$ and $|\mathcal{X}| \geq \log^2 n$, and for all $t \leq \log_{(1+\alpha)\rho}(n^{1-\gamma}/|\mathcal{X}|)$,*

$$\mathbb{P}\left((|\mathcal{N}_t^-(\mathcal{X})|, |\mathcal{N}_t^+(\mathcal{X})|) = (1 \pm \alpha)^t (|\mathcal{X}^-|, |\mathcal{X}^+|) M^t \mid E_H\right) = 1 - o(n^{-2}). \tag{5.76}$$

Proof. For $t \geq 1$ let $d_t^\pm := |\mathcal{N}_t^\pm(\mathcal{X})|$. Furthermore, we let E_t^-, E_t^+ respectively denote the events

$$\frac{(1-\alpha)(\mu_{1,1} d_{t-1}^- + \mu_{2,0} d_{t-1}^+)}{\lambda} \leq d_t^- \leq \frac{(1+\alpha)(\mu_{1,1} d_{t-1}^- + \mu_{2,0} d_{t-1}^+)}{\lambda},$$
$$\frac{(1-\alpha)(\mu_{0,2} d_{t-1}^- + \mu_{1,1} d_{t-1}^+)}{\lambda} \leq d_t^+ \leq \frac{(1+\alpha)(\mu_{0,2} d_{t-1}^- + \mu_{1,1} d_{t-1}^+)}{\lambda}.$$

Define $E_t := E_t^- \cap E_t^+$ and note that the desired event is implied by $\cap_{j=1}^t E_j$. Thus, if we show that $\mathbb{P}\left(E_t^c \mid E_H \cap \left[\cap_{j=1}^{t-1} E_j\right]\right) = o(n^{-3})$, a union bound over all t concludes the proof of (5.76).

If $|\mathcal{X}| \geq n^{1-\gamma}$ there is nothing to prove. Otherwise, the event $\cap_{j=1}^{t-1} E_j$ implies that $d_{t-1}^\pm \leq n^{1-\gamma}$ and $\sum_{j=0}^{t-1}(d_j^- + d_j^+) + |\mathcal{V}(H)| = O(n^{1-\gamma})$.

We run the exploration process described in Section 5.3 starting from $\mathcal{X}$ in order to show that the neighbourhoods of $\mathcal{X}$ have appropriate sizes. We make the following modifications to it:

- Start the process with $\mathcal{A}_0^\pm := \mathcal{X}^\pm$.
- In (iii) if $e_i^* \in \mathcal{L}^*$, we let $\mathcal{A}_i^\pm = \mathcal{A}_i^\pm \backslash \{e_i^\pm\}$.
- In (iii) if $e_i^* \in \mathcal{F}^*(H)$, rather than terminate the process, we let $\mathcal{A}_i^\pm = \mathcal{A}_i^\pm \backslash \{e_i^\pm\}$ and proceed.

This adapts the process to generate a collection of trees rooted in the set $\mathcal{X}$. Note that the union of the t-th level of all such trees comprises $\mathcal{N}_t^+(\mathcal{X}) \cup \mathcal{N}_t^-(\mathcal{X})$. Moreover, a vertex v of a tree T is in $\mathcal{N}_t^+(\mathcal{X})$ if its unique neighbour u in the $(t-1)$-st level of the

tree is such that $\vec{uv}$ is an edge in $\mathbb{DCM}_n$. Similarly it is in $\mathcal{N}_t^-(\mathcal{X})$ if this edge is instead $\vec{vu}$.

As we run the BFS, we can split the exploration into epochs. Recall that i_t is the last step of the t-th epoch i.e., the first time at which all of the stubs at distance at most t from $\mathcal{X}$ have been paired. Note that $i_0 = 0$ and that the number of stubs activated in the t-th epoch is $d_{t+1}^- + d_{t+1}^+$. Let $X_i^\pm = |\mathcal{A}_i^\pm| - |\mathcal{A}_{i-1}^\pm| + 1$. The only way in which we can activate new stubs in the i-th step is if $f_i^\sharp \in \mathcal{U}_{i-1}^\sharp \backslash \mathcal{L}^\sharp$. In this case we activate all the stubs in $\mathcal{E}^\pm(v_i)$ except $f_i^\sharp$. Thus,

$$
\begin{aligned}
X_i^* &= |\mathcal{E}^*(v_i)| \mathbb{1}(f_i^\sharp \in \mathcal{U}_{i-1}^\sharp \setminus \mathcal{L}^\sharp) \\
X_i^\sharp &= (|\mathcal{E}^\sharp(v_i)| - 1) \mathbb{1}(f_i^\sharp \in \mathcal{U}_{i-1}^\sharp \setminus \mathcal{L}^\sharp) \ .
\end{aligned}
$$

Let $\mathcal{H}_{i-1}$ be a history of the process under which $E_H \cap \bigcap_{j=1}^{t-1} E_j$ holds (assuming that i is in the t-th epoch). Suppose that $* = +$ (that is, we match a tail in the $(t-1)$-th level with a head), then for all $(k, \ell) \in \{0, 1, \ldots, K-1\}^2$,

$$
\begin{aligned}
\mathbb{P}\left((X_i^-, X_i^+) = (k-1, \ell) \mid \mathcal{H}_{i-1}\right) &= \frac{\sum_{f^- \in \mathcal{U}_{i-1}^- \backslash \mathcal{L}^-} \mathbb{1}(d_{v(f^-)} = (k, \ell))}{m_n - |\mathcal{P}^-(H)| - (i-1)} \\
&\geq \frac{\sum_{f^- \in \mathcal{E}^- \backslash \mathcal{L}^-} \mathbb{1}(d_{v(f^-)} = (k, \ell))}{m_n} - \frac{|\mathcal{E}^- \backslash \mathcal{U}_{i-1}^-|}{m_n} \\
&\geq \max\{\mathbb{P}\left((D_n)_{\text{in}} = (k-1, \ell)\right) - n^{-\gamma/4}, 0\} \\
&:= b_{n,(k,\ell)}^- \ ,
\end{aligned}
$$

where $(D_n)_{\text{in}}$ is defined as in (5.48). Similary, if $* = -$, we have

$$
\mathbb{P}\left((X_i^-, X_i^+) = (k, \ell-1) \mid \mathcal{H}_{i-1}\right) \geq \max\{\mathbb{P}\left((D_n)_{\text{out}} = (k, \ell-1)\right) - n^{-\gamma/4}, 0\}.
$$

For the sake of simplicity, below we will assume that at time i we have $* = +$. Write $b^- = \sum_{i,j=0}^{K-1} b_{n,(i,j)}^-$ and note $b^- \in [0, 1]$. Let $X_i^\downarrow$ and $X_i^\uparrow$ be independent random variables

on $\mathbb{Z}^2_{\geq 0}$ with distributions

$$\mathbb{P}\left(X_i^{\downarrow} = (k-1,\ell)\right) = \begin{cases} 1 - b^- + b^-_{n,(1,0)}, & \text{if } (k,\ell) = (1,0) \\ b^-_{n,(k,\ell)}, & \text{if } 1 \leq k,\ell \leq K-1 \\ 0 & \text{otherwise} \end{cases}, \tag{5.77}$$

and

$$\mathbb{P}\left(X_i^{\uparrow} = (k-1,\ell)\right) = \begin{cases} b^-_{n,(k,\ell)}, & \text{if } 0 \leq k,\ell \leq K-1 \text{ and } (k,\ell) \neq (K-1,K-1) \\ 1 - b^- + b^-_{n,(K-1,K-1)}, & \text{if } (k,\ell) = (K-1,K-1) \\ 0 & \text{otherwise} \end{cases}. \tag{5.78}$$

Note that $X_i^{\downarrow} \leq (X_i \mid \mathcal{H}_{i-1}) \leq X_i^{\uparrow}$. Furthermore, the mean vectors satisfy

$$\mathbb{E}(X_i^{\downarrow}) = (1+o(1))\mathbb{E}(D_{\text{in}}\mathbb{1}(D^-_{\text{in}} < K,\ D^+_{\text{in}} < K)) - K^2 n^{-\gamma/4} \geq \left(1 - \frac{\alpha}{2}\right)\mathbb{E}(D_{\text{in}}) = \left(1 - \frac{\alpha}{2}\right)(\mu_{2,0}, \mu_{1,1}),$$
$$\mathbb{E}(X_i^{\uparrow}) = (1+o(1))\mathbb{E}(D_{\text{in}}\mathbb{1}(D^-_{\text{in}} < K,\ D^+_{\text{in}} < K)) - K^2 n^{-\gamma/4} \leq \left(1 + \frac{\alpha}{2}\right)\mathbb{E}(D_{\text{in}}) = \left(1 + \frac{\alpha}{2}\right)(\mu_{2,0}, \mu_{1,1}),$$

where the final inequalities follow from (5.74). The analysis with $* = -$ is identical, replacing $(k-1,\ell)$ by $(k,\ell-1)$ and $(\mu_{2,0},\mu_{1,1})$ by $(\mu_{1,1},\mu_{0,2})$; we omit it.

As $X_i^{\downarrow}$ and $X_i^{\uparrow}$ are bounded random variables we may apply Hoeffding's inequality. Furthermore, note that there are d^-_{t-1} and d^+_{t-1} stubs in the t-th epoch such that $* = -$ and $* = +$, respectively. Thus,

$$\begin{aligned}\mathbb{P}\left(d^-_t < (1-\alpha)\frac{\mu_{1,1}d^-_{t-1} + \mu_{2,0}d^+_{t-1}}{\lambda} \mid \mathcal{H}_{i_{t-1}}\right) &\leq \mathbb{P}\left(\sum_{i=i_{t-1}+1}^{i_t} \left((X_i^{\downarrow})^- - \mathbb{E}((X_i^{\downarrow})^-)\right) > \frac{\alpha}{2}\frac{\mu_{1,1}d^-_{t-1} + \mu_{2,0}d^+_{t-1}}{\lambda}\right. \\ &\leq \exp\left(-\frac{\alpha^2(\mu_{1,1}d^-_{t-1} + \mu_{2,0}d^+_{t-1})^2}{8K^2\lambda^2(d^-_{t-1} + d^+_{t-1})}\right) = o(n^{-3})\end{aligned}$$

as $d^-_{t-1} + d^+_{t-1} \geq |\mathcal{X}| \geq \log^2 n$ by $\mathcal{H}_{i_{t-1}}$. The analogous results also hold for d^+_t. Moreover,

in the same way by considering $(X_i^{\uparrow})^-$,

$$\mathbb{P}\left(d_t^- > (1+\alpha)\frac{\mu_{1,1}d_{t-1}^- + \mu_{2,0}d_{t-1}^+}{\lambda} \mid \mathcal{H}_{i_{t-1}}\right) = o(n^{-3})\,.$$

and similarly for d_t^+. □

Proof of Proposition 5.5.1. Let $\alpha, \beta > 0$ and define $t^{\mathcal{X}} := \lceil \log_{(1-\alpha)\rho}(n^{1/2+\beta}/|\mathcal{X}|)\rceil$ as well as $t^{\mathcal{Y}} := \lceil \log_{(1-\alpha)\rho}(n^{1/2+\beta}/|\mathcal{Y}|)\rceil$. If α and β are sufficiently small with respect to ε, then $t^{\mathcal{X}} + t^{\mathcal{Y}} + 1 \leq (1+\varepsilon)\log_\rho n$.

If a stub in $\mathcal{N}^{\pm}_{<t^{\mathcal{X}}}(\mathcal{X})$ is paired to a stub in $\mathcal{N}^{\pm}_{<t^{\mathcal{Y}}}(\mathcal{Y})$, then we are done and $\mathrm{dist}(\mathcal{X},\mathcal{Y}) \leq t^{\mathcal{X}} + t^{\mathcal{Y}}$. Let assume otherwise and recall that the matrix M is positive and has leading eigenvalue ρ. Hence, by the Perron-Frobenius theorem, $\rho^{-t}M^t \to \mathbf{u}^\top\mathbf{v}$ where $\mathbf{u}$ and $\mathbf{v}$ are the dominant right and left eigenvectors of M respectively (chosen such that $\mathbf{v}\mathbf{u}^\top = 1$).

For $\delta > 0$ which we shall take to be arbitrarily small, let $B^\delta = \mathbf{u}^\top\mathbf{v} - \delta J$, where J is the all 1s matrix. By Lemma 5.5.2 and the choice of $t^{\mathcal{X}}$, for any $\delta > 0$ it follows that with probability at least $1 - o(n^{-2})$

$$(|\mathcal{N}^-_{t^{\mathcal{X}}}(\mathcal{X})|, |\mathcal{N}^+_{t^{\mathcal{X}}}(\mathcal{X})|) \geq (1-\alpha)^{t^{\mathcal{X}}}(|\mathcal{X}^-|, |\mathcal{X}^+|)M^{t^{\mathcal{X}}} \geq n^{\frac{1}{2}+\beta}\left(\frac{|\mathcal{X}^-|}{|\mathcal{X}|}, \frac{|\mathcal{X}^+|}{|\mathcal{X}|}\right)B^\delta =: (c^-_{\mathcal{X}}, c^+_{\mathcal{X}})n^{\frac{1}{2}+\beta} \tag{5.79}$$

Similarly, define $c^{\pm}_{\mathcal{Y}}$ for the set $\mathcal{Y}$. The probability that there is no pairing between $\mathcal{N}^{\pm}_{t^{\mathcal{X}}}(\mathcal{X})$ and $\mathcal{N}^{\pm}_{t^{\mathcal{Y}}}(\mathcal{Y})$ is at most

$$\left(1 - \frac{c^-_{\mathcal{X}}n^{1/2+\beta}}{m_n}\right)^{c^+_{\mathcal{Y}}n^{1/2+\beta}}\left(1 - \frac{c^+_{\mathcal{X}}n^{1/2+\beta}}{m_n}\right)^{c^-_{\mathcal{Y}}n^{1/2+\beta}} = o(n^{-3}).$$

Therefore the probability that $\mathrm{dist}(\mathcal{X},\mathcal{Y}) > t^{\mathcal{X}} + t^{\mathcal{Y}} + 1$ is $o(n^{-3})$ which proves (5.73).

□

5.6 Proof of Theorem 5.1.2

In this section we conclude the proof of our main theorem.

5.6.1 Supercritical case

Suppose that $\rho > 1$, here we will prove (5.8). We first show that most of the vertices in $\mathcal{L}_V$ belong to the same WCC. Fix ε and recall the definition of $\mathcal{L}_V(\varepsilon)$ in (5.68). Let $x, y \in \mathcal{L}_V(\varepsilon)$, then there exists $t_x, t_y \leq (1+\varepsilon)t_0$ such that both $\mathcal{X} = \mathcal{N}^{\pm}_{t_x}(x)$ and $\mathcal{Y} = \mathcal{N}^{\pm}_{t_y}(y)$ have size at least ω. Moreover, the graph H induced by $\mathcal{N}^{\pm}_{\leq t_x}(x) \cup \mathcal{N}^{\pm}_{\leq t_y}(y)$ satisfies $|\mathcal{V}(H)| \leq 4(1+\varepsilon)\omega t_0 = O(\log^7 n)$. So we may apply Proposition 5.5.1, to conclude that there exists a sequence of edges connecting $\mathcal{X}$ and $\mathcal{Y}$. It follows that x and y belong to the same WCC, and by the proof of Proposition 5.4.4, whp this component has order at least $|\mathcal{L}_V(\varepsilon)| = (1+o(1))\eta n$ and size at least $|\mathcal{L}_E(\varepsilon)| = (1+o(1))\zeta n$.

We now bound from above the order and size of the WCC that contains $\mathcal{L}_V(\varepsilon)$. Fix $\delta > 0$. For $K \in \mathbb{N}$, let S_K be the set of vertices with either in-degree or out-degree at least K. Fix K large enough so that $\sum_{v \in S_K}(d_v^+ + d_v^-) \leq \delta n$. The existence of such K is guaranteed by Condition 5.1.1. ;in particular, by the convergence of the second moments we can choose $K = o(\delta^{-1/2})$. Observe that $|S_K| \leq \delta n$.

Choose $v \notin \mathcal{L}_V \cup S_K$ and let us compute the probability that v belongs to the same WCC as $\mathcal{L}_V(\varepsilon)$. If it does, as all vertices in $\mathcal{L}_V(\varepsilon)$ appear in a neighbourhood of v, we have that $|\mathcal{N}_t^-(v)|, |\mathcal{N}_t^+(v)| \in (0, \omega)$ for all $t \leq |\mathcal{L}_V(\varepsilon)|/2\omega$. By Lemma 5.4.3 the probability of this event is at most $\hat{\rho}^{|\mathcal{L}_V(\varepsilon)|/2\omega - O(\log\log n)} = o(n^{-1})$. By a union bound over all vertices, we have that whp, no such vertex v exists. Thus, there are at most $|S_K| \leq \delta n$ vertices not in $\mathcal{L}_V$ that belong to the WCC containing $\mathcal{L}_V(\varepsilon)$. It follows that this component has order at most $|\mathcal{L}_V| + \delta n$, which is at most $(\eta + 2\delta)n$ whp. Similarly, the size of the WCC is at most $(\zeta + \delta + 2K\delta)n = (\zeta + o(\delta^{1/2}))n$. Since δ can be made arbitrarily small, we conclude the proof of (5.8).

Following the same arguments, the previous proof shows that any other component

has size at most $\log^c n$, for some constant $c > 0$.

5.6.2 Subcritical Case

We finish by proving (5.10). We shall consider the order of all of the components of the random digraph. Define $\mathcal{C}(v)$ to be the component of $\mathbb{DCM}$ which contains v. The result is implied by the following,

Lemma 5.6.1. *If $\rho < 1$ then whp*

$$\sum_{v \in [n]} |\mathcal{C}(v)| = o(n^2). \tag{5.80}$$

Proof of (5.10). Suppose for a contradiction that there is a component $\mathcal{C}_0$ with $|\mathcal{C}_0| \geq \varepsilon n$. Then, $\sum_{v \in [n]} |\mathcal{C}(v)| \geq |\mathcal{C}_0|^2 \geq \varepsilon^2 n^2$ contradicting Lemma 5.6.1. □

Proof of Lemma 5.6.1. Fix a vertex $v \in [n]$ and denote by $\mathcal{N}_t = \mathcal{N}_t^-(v) \cup \mathcal{N}_t^+(v)$ its t-th neighbourhood. Let $\beta > 0$ be sufficiently small. We shall call v *big* if any of the following is true and *small* otherwise:

i) $d_v^- + d_v^+ \geq n^{\beta/3}$;

ii) $|\mathcal{N}_t| \geq n^{\beta/2}$ for some $t \in \mathbb{N}$;

iii) $|\mathcal{N}_{t_0}| \geq 1$ for $t_0 = \frac{\log n}{\log(1/\rho)}$.

All small vertices are contained in components of order at most $K = n^{\beta/2} h_0 = o(n)$.

We bound the probability of each way in which v can be big in turn. To bound i), let I be a uniformly random element of $[n]$. Using the bounded second moment in Condition 5.1.1, we have

$$\mathbb{P}\left(d_I^- + d_I^+ \geq n^{\beta/3}\right) \leq \mathbb{P}\left(d_I^- \geq n^{\beta/3}/2\right) + \mathbb{P}\left(d_I^+ \geq n^{\beta/3}/2\right) = O(n^{-2\beta/3}) \,. \tag{5.81}$$

So we may assume that $d_v^- + d_v^+ \leq n^{\beta/3}$ and bound ii) and iii) using branching processes. Let $(X_t)_{t\geq 1}$ be the total size of the t-th generation of $\mathrm{GW}^{(d_v^-,d_v^+)}_{\Xi_n^\uparrow(\beta)}$, as defined in Subsection 5.3.2. Determining whether v is big or small requires revealing at most $K = O(n^{\beta/2}\log n)$ half-edges. Therefore, we can use the coupling of neighbourhoods with multitype Galton-Watson trees in Lemma 5.3.1 to compute the probability v is big, and we will study the latter.

Write $\mathbf{d} = (d_v^-, d_v^+)$. Recall the definition of the mean matrix M in (5.6) with largest eigenvalue ρ. Choose $\varepsilon > 0$ small enough. The Frobenius norm of the t-th power satisfies $\|M^t\| \leq C\rho^{(1-\varepsilon)t}$ for some $C = C(\varepsilon)$. We have

$$\mathbb{E}(X_t) = \mathbf{d}M^t\mathbf{1}^\top \leq \sqrt{2}\|\mathbf{d}\|\|M^t\| \leq \sqrt{2}Cn^{\beta/3}\rho^{(1-\varepsilon)t}. \tag{5.82}$$

By Markov's inequality and union bound over $t \geq 1$, we bound the probability of ii):

$$\mathbb{P}\left(\exists t:\, X_t \geq n^{\beta/2}\right) \leq \sum_{t\geq 1}\mathbb{P}\left(X_t \geq n^{\beta/2}\right) \leq \sqrt{2}Cn^{-\beta/6}\sum_{t\geq 1}\rho^{(1-\varepsilon)t} = O(n^{-\beta/6}). \tag{5.83}$$

Finally, we bound the probability of iii). By Markov's inequality,

$$\mathbb{P}\left(X_{t_0} \geq 1\right) \leq \sqrt{2}Cn^{\beta/3}\rho^{(1-\varepsilon)t_0} = O(n^{\varepsilon+\beta/3-1}). \tag{5.84}$$

Combining (5.81), (5.83) and (5.84), we deduce that the probability a vertex is big is $o(1)$.

We thus have

$$\mathbb{E}\left(\sum_{v\in[n]}|\mathcal{C}(v)|\right) \leq \mathbb{E}\left(\sum_{v\in[n]}\Big(|\mathcal{C}(v)|\mathbb{1}(v \text{ small}) + |\mathcal{C}(v)|\mathbb{1}(v \text{ big})\Big)\right) \leq nK + o(1)n^2 = o(n^2).$$

So (5.80) holds with high probability by Markov's inequality. □

CHAPTER 6

CONCLUDING REMARKS

In this thesis we have proven a number of new results on thresholds for the giant component in various models of random graphs and digraphs. This contributes to the picture of what the threshold looks like however there are many questions that still remain in all of the models which we have studied and we describe some of these questions here.

Firstly, in $D(n,p)$ we have proved tail bounds on the size of the largest component in the critical window. The recent result of Goldschmidt and Stephenson [31] complements our result by also proving that there is a scaling limit for the rescaled size of the largest components. A fully explicit description of the random variable X such that the largest component is of order $Xn^{1/3}$ would be a great addition to these two results. Perhaps a result similar to Pittel's [69] in $G(n,p)$ may be possible. Also, recall that the size of the largest component in the barely subcritical case was solved by Łuczak and Seierstad. They also showed that the next k components for any finite k are also of the same order as the first in this regime. As such a natural question to ask is whether there is a scaling limit for the joint distribution of the components in descending order.

In the configuration model for graphs we proved a bound on the size of the largest component in the barely subcritical regime. This is the first result in this regime which is better than $o(n)$. There are certainly improvements which could be made to the assumptions which we use. In particular it would be interesting to know whether it is possible to remove the condition on the fourth moment or if we could increase the

maximum degree for which our result works. Also, in the critical window Dhara et. al. [17] showed that if the third moment of the degree sequence is finite then there is a Brownian motion type scaling limit. It would be interesting to know how far this can be extended to infinite third moment.

In the configuration model for digraphs, we have two results. First, we showed a result on the barely subcritical regime for the giant strongly connected component. We found the size and structure of these components. In an upcoming work [14] we show that there is a complementary result for the barely supercritical regime. The question of exactly what happens inside the critical window is still an open question however. We also found the threshold for a giant weakly connected component confirming a prediction of Kryven [49]. At present this is the only result on the weak component in the directed configuration model. As such the questions of what happens in the barely subcritical, critical window and barely supercritical ranges as well as precisely where these ranges lie are all still open.

LIST OF REFERENCES

[1] L. Addario-Berry, N. Broutin, and C. Goldschmidt. The continuum limit of critical random graphs. *Probability Theory and Related Fields*, 152(3-4):367–406, 2012.

[2] D. Aldous. Brownian excursions, critical random graphs and the multiplicative coalescent. *The Annals of Probability*, pages 812–854, 1997.

[3] K.B. Athreya and P. Ney. *Branching Processes.* Springer-Verlag, Berlin Heidelberg, 1972.

[4] S. Balogh, G. Palla, and I. Kryven. Networks with degree-degree correlations are special cases of the edge-coloured random graph. *Journal of Complex Networks*, 8(4), 2020.

[5] J. Blanchet and A. Stauffer. Characterizing optimal sampling of binary contingency tables via the configuration model. *Random Structures & Algorithms*, 42(2):159–184, 2013.

[6] B. Bollobás. A probabilistic proof of an asymptotic formula for the number of labelled regular graphs. *European Journal of Combinatorics*, 1(4):311–316, 1980.

[7] B. Bollobás. The evolution of random graphs. *Transactions of the American Mathematical Society*, 286(1):257–274, 1984.

[8] B. Bollobás and O. Riordan. An old approach to the giant component problem. *J. Combin. Theory (Series B)*, 113:236–260, 2015.

[9] C. Borgs, J. Chayes, S. Dhara, S. Petti, and S. Sen. Sampling from sparse graphs with overlapping communities and heterogeneous degrees. 2021.

[10] C. Butts, M. Petrescu-Prahova, and B. Remy Cross. Responder communication networks in the World Trade Center disaster: Implications for modeling of communication within emergency settings. *Mathematical Sociology*, 31(2):121–147, 2007.

[11] X.S. Cai and G. Perarnau. The giant component of the directed configuration model revisited. *arXiv preprint arXiv:2004.04998*, 2020.

[12] J. Cao and M. Olvera-Cravioto. Connectivity of a general class of inhomogeneous random digraphs. *Random Structures & Algorithms*, 56(3):722–774, 2020.

[13] C. Cooper and A. Frieze. The size of the largest strongly connected component of a random digraph with a given degree sequence. *Combinatorics, Probability and Computing*, 13(3):319–337, 2004.

[14] M. Coulson. Barely supercritical random digraphs with a given degree sequence. *Manuscript*.

[15] M. Coulson. The critical window in random digraphs. *Combinatorics, Probability and Computing*, pages 1–19, 2021.

[16] M. Coulson and G. Perarnau. Weak components of the directed configuration model. In *Extended Abstracts EuroComb 2021*, pages 682–687. Springer International Publishing, 2021.

[17] S. Dhara, R. van der Hofstad, J. van Leeuwaarden, and S. Sen. Critical window for the configuration model: finite third moment degrees. *Electronic Journal of Probability*, 22, 2017.

[18] S. Dhara, R. van der Hofstad, J. van Leeuwaarden, and S. Sen. Heavy-tailed configuration models at criticality. In *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, volume 56, pages 1515–1558. Institut Henri Poincaré, 2020.

[19] S. Donderwinkel and Z. Xie. Universality for the directed configuration model with random degrees: metric space convergence of the strongly connected components at criticality. *arXiv preprint arXiv:2105.11434*, 2021.

[20] R. Doney. One-sided local large deviation and renewal theorems in the case of infinite mean. *Probability theory and related fields*, 107(4):451–465, 1997.

[21] R. Durrett. *Random graph dynamics*, volume 200. Cambridge University Press, 2007.

[22] R. Durrett. *Probability: theory and examples*, volume 49. Cambridge university press, 2019.

[23] P. Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[24] P. Erdos and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[25] T. Erhardsson. Steins method for Poisson and compound Poisson approximation. *An introduction to Stein's method*, 4:61–114, 2005.

[26] S. Eubank, H. Guclu, A. Kumar, M. Marathe, A. Srinivasan, Z. Toroczkai, and N. Wang. Modelling disease outbreaks in realistic urban social networks. *Nature*, 429(6988):180–184, 2004.

[27] N. Fountoulakis. Percolation on sparse random graphs with given degree sequence. *Internet Mathematics*, 4(4):329–356, 2007.

[28] N. Fountoulakis and B. Reed. A general critical condition for the emergence of a giant component in random graphs with given degrees. *Electronic Notes in Discrete Mathematics*, 34:639–645, 2009.

[29] D.A. Freedman. On tail probabilities for martingales. *The Annals of Probability*, pages 100–118, 1975.

[30] A. Frieze and M. Karoński. *Introduction to random graphs*. Cambridge University Press, 2016.

[31] C. Goldschmidt and R. Stephenson. The scaling limit of a critical random directed graph. *arXiv preprint arXiv:1905.05397*, 2019.

[32] A. Graf. On the strongly connected components of random directed graphs with given degree sequences. Master's thesis, University of Waterloo, 2016.

[33] G. Grimmett and D. Stirzaker. *Probability and random processes*. Oxford University Press, 2001.

[34] T.E. Harris. *The Theory of Branching Processes*. Springer-Verlag, Berlin Heidelberg, 1963.

[35] H. Hatami and M. Molloy. The scaling window for a random graph with a given degree sequence. *Random Structures & Algorithms*, 41(1):99–123, 2012.

[36] R. van der Hofstad. *Random Graphs and Complex Networks*, volume 1 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, England, 2016.

[37] R. van der Hofstad, S. Janson, and M. Luczak. Component structure of the configuration model: barely supercritical case. *Random Structures & Algorithms*, 55(1):3–55, 2019.

[38] F. Hoppe. Supercritical multitype branching processes. *The Annals of Probability*, 4(3):393–401, 1976.

[39] P. Jagers and A. Lagerås. General branching processes conditioned on extinction are still branching processes. *Electronic Communications in Probability*, 13:540–547, 2008.

[40] S. Janson. The largest component in a subcritical random graph with a power law degree distribution. *The Annals of Applied Probability*, 18(4):1651–1668, 2008.

[41] S. Janson. The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225, 2009.

[42] S. Janson, D.E. Knuth, T. Łuczak, and B. Pittel. The birth of the giant component. *Random Structures & Algorithms*, 4(3):233–358, 1993.

[43] S. Janson and M. J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.

[44] S. Janson, T. Łuczak, and A. Ruciński. *Random graphs*, volume 45. John Wiley & Sons, 2011.

[45] F. Joos, G. Perarnau, D. Rautenbach, and B. Reed. How to determine if a random graph with a fixed degree sequence has a giant component. *Probability Theory and Related Fields*, 170(1-2):263–310, 2018.

[46] A. Joseph. The component sizes of a critical random graph with pre-described degree sequence. *The Annals of Applied Probability*, 24:2560–2594, 2014.

[47] R.M. Karp. The transitive closure of a random digraph. *Random Structures & Algorithms*, 1(1):73–93, 1990.

[48] S.M. Kirsch, M. Gnasa, and A.B. Cremers. Beyond the Web: Retrieval in Social Information Spaces. In *Advances in Information Retrieval*, pages 84–95, Berlin, Heidelberg, 2006. Springer, Springer.

[49] I. Kryven. Emergence of the giant weak component in directed random graphs with arbitrary degree distributions. *Phys Rev E*, 94(1):012315, 2016.

[50] I. Kryven. Bond percolation in coloured and multiplex networks. *Nature communications*, 10(1):1–16, 2019.

[51] L. Lu and L. Székely. Using Lovász Local Lemma in the space of random injections. *the electronic journal of combinatorics*, page R63, 2007.

[52] T. Łuczak. Component behavior near the critical point of the random graph process. *Random Structures & Algorithms*, 1(3):287–310, 1990.

[53] T. Łuczak. The phase transition in the evolution of random digraphs. *Journal of graph theory*, 14(2):217–223, 1990.

[54] T. Łuczak, B. Pittel, and J.C. Wierman. The structure of a random graph at the point of the phase transition. *Transactions of the American Mathematical Society*, 341(2):721–748, 1994.

[55] T. Łuczak and T. G. Seierstad. The critical behavior of random digraphs. *Random Structures & Algorithms*, 35(3):271–293, 2009.

[56] A. Martin-Löf. Symmetric sampling procedures, general epidemic processes and their theshold limit theorems. *Journal of Applied Probability*, 23(2):265–282, 1986.

[57] C. McDiarmid. Clutter percolation and random graphs. In *Combinatorial Optimization II*, pages 17–25. Springer, 1980.

[58] M. Molloy and B. Reed. A critical point for random graphs with a given degree sequence. *Random Structures Algorithms*, 6:161–180, 1995.

[59] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Comb., Probab. Comput.*, 7:295–305, 1998.

[60] A.B. Mukhin. Local limit theorems for distributions of sums of independent random vectors. *Theory of Probability & Its Applications*, 29(2):369–375, 1985.

[61] A.B. Mukhin. Local limit theorems for lattice random variables. *Theory of Probability & Its Applications*, 36(4):698–713, 1992.

[62] A. Nachmias and Y. Peres. Critical percolation on random regular graphs. *Random Structures & Algorithms*, 36(2):111–148, 2010.

[63] A. Nachmias and Y. Peres. The critical random graph, with martingales. *Israel Journal of Mathematics*, 176(1):29–41, 2010.

[64] M.E.J Newman, S.H. Strogatz, and D.J. Watts. Random graphs with arbitrary degree distributions and their applications. *Physical review E*, 64(2):026118, 2001.

[65] I. Palásti. On the strong connectedness of directed random graphs. *Studia Sci. Math. Hungar*, 1:205–214, 1966.

[66] X. Pérez-Giménez and N. Wormald. Asymptotic enumeration of strongly connected digraphs by vertices and edges. *Random Structures & Algorithms*, 43(1):80–114, 2013.

[67] V.V. Petrov. *Sums of independent random variables.* Springer-Verlag, New York-Heidelberg, 1975. Translated from the Russian by A. A. Brown, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 82.

[68] S. Petti. *Randomness as a Tool for Modeling and Uncovering Structure.* PhD thesis, Georgia Institute of Technology, 2020.

[69] B. Pittel. On the largest component of the random graph at a nearcritical stage. *Journal of Combinatorial Theory, Series B*, 82(2):237–269, 2001.

[70] B. Pittel. On the largest component of a random graph with a subpower-law degree sequence in a subcritical phase. *The Annals of Applied Probability*, pages 1636–1650, 2008.

[71] B. Pittel. Counting strongly-connected, moderately sparse directed graphs. *Random Structures & Algorithms*, 43(1):49–79, 2013.

[72] O. Riordan. The phase transition in the configuration model. *Comb., Probab. Comput.*, 21:265–299, 2012.

[73] A.D. Scott and G.B. Sorkin. Solving sparse random instances of max cut and max 2-csp in linear expected time. *Combinatorics, Probability and Computing*, 15(1-2):281–315, 2006.

[74] I.S. Tyurin. Refinement of the upper bounds of the constants in Lyapunov's theorem. *Russian Mathematical Surveys*, 65(3):586–588, 2010.

[75] S.S. Venkatesh. *The theory of probability: Explorations and applications.* Cambridge University Press, 2013.

[76] H.M. Wang. On total progeny of multitype Galton-Watson process and the first passage time of random walk on lattice. *Acta Mathematica Sinica, English Series*, 30(12):2161–2172, 2014.

[77] D. Williams. *Probability with martingales.* Cambridge university press, 1991.

[78] E.M. Wright. Formulae for the number of sparsely-edged strong labelled digraphs. *The Quarterly Journal of Mathematics*, 28(3):363–367, 1977.